Université Paris 6 Doctoral Dissertation

The Traveling Salesman and Related Stochastic Problems

by

Allon Percus

A dissertation submitted for the degree of
Docteur de l'Université Paris 6
in Theoretical Physics

22 September 1997

PhD Committee:	**Claire Kenyon**	*Reader*
	Olivier Martin	*Thesis advisor*
	Rémi Monasson	
	Karol Penson	*Committee chair*
	Nicolas Sourlas	*Reader*
	Stéphane Zaleski	

Abstract

In the traveling salesman problem, one must find the length of the shortest closed tour visiting given "cities". We study the stochastic version of the problem, taking the locations of cities and the distances separating them to be random variables drawn from an ensemble. We consider first the ensemble where cities are placed in Euclidean space. We investigate how the optimum tour length scales with number of cities and with number of spatial dimensions. We then examine the analytical theory behind the random link ensemble, where distances between cities are independent random variables. Finally, we look at the related geometric issue of nearest neighbor distances, and find some remarkable universalities.

Foreword

Université Paris 6 authorities require PhD candidates to submit and defend their dissertations in French. As I find this policy unhelpful to the scientific community at large, the manuscript that I have decided to present here for public distribution is the original (unofficial) English-language version of the text.

• • •

My graduate student years owe their successful completion — in both a personal and a scientific sense — to a very large number of individuals. In the limited space I have here, let me highlight the contributions of those to whom I am most grateful.

- **Olivier Martin**, my thesis advisor. I imagine that anyone who has ever worked with Olivier has found the experience to be an immense pleasure, from start to finish. I am delighted to have shared in this pleasure. Not only is Olivier an extraordinarily warm, friendly and patient person; he also has one of the liveliest scientific minds I have ever encountered. If my three years of thesis work have been among the happiest of my life, Olivier deserves a great deal of the credit.

- **Dominique Vautherin**, former head of our *Division de Physique Théorique*, whose efforts to facilitate my stay here went far above and beyond the call of duty. His hospitality and constant encouragement played a large part in making my time at Orsay productive and enjoyable.

- Other members of the *Division de Physique Théorique*, who contributed to the fruitful environment of these past three years: **Oriol Bohigas**, whose ideas were fundamental to much of our work; **Xavier Campi**, present head of the Division; **Nadine Rohart**, Division secretary, whose skill at solving the most complex administrative problems in under thirty seconds is nothing short of remarkable; **Michèle Verret**, secretary, who was a constant and invaluable source of help to me, and who willingly took on the task of printing and binding this thesis.

- The students and postdocs in Olivier's group: **Georg Schreiber**, **Jérôme Houdayer**, **Jacques Boutet de Monvel** and **Nicolas Cerf**, who could all be relied upon at any moment for stimulating discussions, friendship and moral support.

- Other students in the Division, who managed to create an atmosphere that was both scientifically productive and socially relaxed: **Alexis Prevost**, **Amaury Mouchet**, **Ulrich Gerland**, **Isabela Porto Cavalcante**, **Christophe Texier**, **Daniel Rouben** and **Guglielmo Iacomelli**. Amaury deserves special mention — the formatting style of his own PhD thesis served as a model for this manuscript, to the point where practically all of the TEX macros used here are the result of his laborious efforts.

- My family, and some of my closest friends: **Marie Aronson**, **Marjorie Bertolus**, **Isabelle Billig**, **Paul Boubli**, **Eleonore Dailly**, **Frédéric Huin**, **Isabelle Kraus**, **Karen Meyer-Roux**, **Anouchka Roggeman**, **Tony Seward**, **Eva Sibony** and **Oana Tataru**. It is difficult for me to express in words how much their constant support and affection has meant to me during these past few years in France.

- My PhD committee, for their dedicated efforts and useful suggestions.

- And last but not least ... those whom I have unintentionally forgotten here, but who surely deserve every bit as much thanks as the individuals I have mentioned. I hope they will accept my apologies and, as this is a first offense, let me off lightly this time. (Next time they will not be so lenient.)

Table of Contents

I Introduction **7**

II Scaling laws in the Euclidean TSP **17**

Chapter overview 17

Reprint: Physical Review Letters 21

Reprint: Journal de Physique 27

III The random link TSP and its analytical solution **49**

Chapter overview 49

Abstract 53

1. Introduction 55
2. Background and the cavity method 56
3. Numerical analysis: $d = 2$ case 60
4. Numerical analysis: renormalized model 64
5. Conclusion 67

Bibliography 68

IV Universalities in nearest neighbor distances **71**

Chapter overview 71

Abstract 73

1. Introduction 75
2. Preliminaries 75
3. Flat Surfaces 77
4. Curved Surfaces 78
5. Regge Calculus 82
6. Conclusions 83

Bibliography 83

Appendix A — Computational complexity: P vs. NP **85**

Appendix B — Outline of a self-averaging proof **87**

Appendix C — Estimates for a non-uniform distribution of cities in space **89**

Appendix D — Numerical methodology **91**

Appendix E — Another numerical study of $\beta(d)$ **97**

General References **101**

Chapter I

Introduction

Consider N sites ("cities") in a space, and the distances connecting each pair of cities. Now find the *shortest* possible path ("tour") visiting each city exactly once, and returning to its starting point. This is the traveling salesman problem (TSP), the most classic and best known problem in the field of combinatorial optimization (CO).

The TSP is an extremely simple problem to state; deceptively so, because it is also an extremely difficult one to solve. The most naïve algorithm for finding the optimum tour would have to consider all $(N-1)!/2$ possible tours. Working this way, the fastest computer on the market today would require more time than the current age of the universe to solve a case with 27 cities. Of course, far more sophisticated algorithms exist, but even the best of these is sharply limited by the inherent complexity of the TSP. Computer scientists classify the problem as being NP-hard: roughly speaking this means that, most likely, there exists no algorithm that can consistently find the optimal tour in an amount of time polynomial in N. The statement provides a technical standard for "hardness", saying that if a polynomial algorithm *could* ever be found for the TSP, one would exist for a whole class of other CO problems as well (Garey & Johnson, 1979, see also discussion in Appendix A). For this reason, and because of how straightforward the TSP is to state, it has served as a prototype for the study of difficult combinatorial optimization problems in general.

The challenge of making progress on the TSP despite its complexity has carried its appeal well beyond its initial boundaries. Research on the TSP and related CO issues has become a truly interdisciplinary effort, uniting computer scientists, pure mathematicians, operations researchers, and more recently, theoretical physicists.[1] In physics, the TSP has been of particular interest in its *stochastic* form, where the configuration ("instance") of the N cities is randomly chosen from an ensemble (see Figure I–1 for an example of an $N = 24$ instance and its optimal tour). Starting in the 1980s, Kirkpatrick, Mézard, Parisi and other theoreticians began applying tools from statistical mechanics to the stochastic TSP, noting broad similarities between CO problems and the emerging field of *disordered systems*. The hope was, first of all, that numerical methods developed in the context of disordered systems could be adapted to the TSP, and second of all, that the analytical formalism developed for these other problems could provide a theoretical framework for understanding the TSP — and ultimately, could provide exact results as well.

The generic example of a disordered system is the spin glass. A spin glass may be thought of as a ferromagnetic Ising model — the fundamental lattice model that statistical physicists use

[1] Biology, in the guise of genetic algorithms and DNA computing, is but the latest field to arrive on the scene.

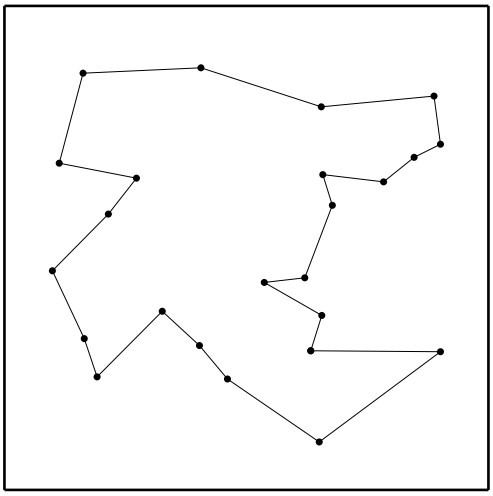

Figure I–1: An elementary instance of the traveling salesman problem, and its optimal tour.

to describe magnetism — but modified so that the interaction strengths between neighboring lattice sites are random quantities. If, as in the case of the spin glass, they are allowed to be negative as well as positive (antiferromagnetic as well as ferromagnetic), we have a *frustrated* disordered system. Take a fixed instance of these random variables ("quenched disorder"). The total energy of the system, involving a sum over all pairs of neighboring sites, is then a complicated function of the particular state of the system (meaning the binary values of the *spins* located at each of N sites), and may be represented schematically by a "landscape" of hills and valleys (see Figure I–2). We shall concern ourselves with what statistical physicists think of as the *zero-temperature* case, where we must find the lowest-energy state, and hence the global minimum in this landscape. If we are to avoid actually enumerating all possible states — a task which increases exponentially with the number of sites — the problem of finding the global minimum becomes highly nontrivial since there are numerous local minima in which a minimizing search procedure could get trapped.

Let us now examine the TSP using this picture. The system's "energy" is the tour length, and is to be minimized. The quenched "interaction strengths" are the distances between cities. The energy is once again a complicated function of the state of the system, where in this case "state" refers to the particular tour chosen. Optimizing the tour is precisely equivalent to finding the system's ground state, and thus the global minimum in Figure I–2. If we look at it in this way, the same approximation methods used to find spin glass ground states may be used to find optimal TSP tours; simulated annealing, developed independently by KIRKPATRICK (1984) and ČERNY (1985), is the most famous example. At the same time, if we look not at a single instance but rather at the stochastic ensemble of instances, analytical approaches giving exact results for spin glass problems can be adapted to provide predictions for the TSP that avoid algorithmic procedures altogether.

The foregoing analogy can be broadened: disordered systems and CO problems display many similar properties, often making methods developed in one field applicable to the other. A common feature in combinatorial optimization is the existence of a large number of quantities interacting in a complicated manner, where each quantity exerts an influence on the others — and these influences are often contradictory. Describing such systems on a microscopic level is generally so difficult as to be impossible. A method better suited to these problems is that of statistical physics, where one contents oneself with studying the broader picture, reducing the number of parameters considerably by resorting to a probabilistic approach. One thus obtains the macroscopic quantities of interest, to reasonable accuracy, with far less effort.

Let us express these concepts more explicitly. A CO problem such as the TSP may be formulated as follows. We have an objective function representing "cost" — the tour length, for the TSP — that must be minimized. This cost may be expressed as

$$L\left(\{n_{ij}\}\right) = \sum_{i,j} l_{ij} n_{ij}, \tag{I.1}$$

where i and j go from 1 up to N (the total system "size"), the $\{l_{ij}\}$ are the parameters of the problem, and the $\{n_{ij}\}$ are the variables whose values must be adjusted so as to minimize L. A certain set of constraints determines the allowed values for the $\{n_{ij}\}$. In the case of the TSP, for example, L is simply the tour length to be minimized, $\{l_{ij}\}$ is the matrix of distances between city i and city j, and $\{n_{ij}\}$ is a binary matrix expressing whether the link between i and j is used on the tour. The constraints on the variables are thus that: (i) each link is either used once or not at all ($\forall i, j,\ n_{ij} \in \{0, 1\}$ and $\forall i,\ n_{ii} = 0$); (ii) each city has one link going into it and one link coming out from it ($\forall j,\ \sum_i n_{ij} = 1$ and $\forall i,\ \sum_j n_{ij} = 1$); and (iii) the $\{n_{ij}\}$ are such that

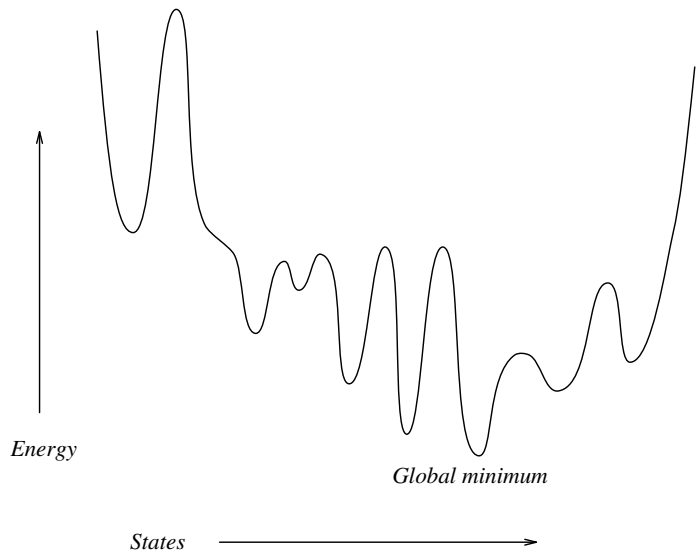

Figure I–2: Schematic energy "landscape", over state space, of a frustrated disordered system.

the links form exactly one loop ($\forall i_1,\ \exists i_2, \ldots, i_N\ : n_{i_1 i_2} = n_{i_2 i_3} = \cdots = n_{i_{N-1} i_N} = n_{i_N i_1} = 1$).[2]

To put this into the language of statistical physics, we think of L as the system's energy; each contribution $l_{ij} n_{ij}$ is then the interaction energy associated with the pair of sites i and j. We may write down the partition function:

$$Z = \sum_{\{n_{ij}\}} e^{-L(\{n_{ij}\})/T}, \tag{I.2}$$

where the sum is taken over the ensemble of matrices $\{n_{ij}\}$ that satisfy the constraints we have posed, and T is the physical "temperature" of the system. In the limit $T \to 0$ the free energy $F = -T \ln Z$ will be dominated by the ground-state contribution, that is, the particular state (value of the matrix $\{n_{ij}\}$) that minimizes L. When we speak of the $T = 0$ solution, we are thus speaking of the global optimum.

Rephrasing the problem in this manner does not appear to help us find the exact state minimizing L for a given set of parameters $\{l_{ij}\}$ — after all, when N is large and we are dealing with matrices of size $N \times N$, there is still an enormous number of parameters. However, we may instead consider the $\{l_{ij}\}$ as being chosen randomly from an ensemble, and ask for properties of the *distribution* of L, such as its mean over the ensemble. Analytical tools from statistical physics, and particularly the approximations that are typically used, then become relevant and useful. The stochastic TSP is a prime candidate for this sort of treatment.

Two main versions of the stochastic TSP have been discussed in the literature, and form the subject of the majority of this thesis. The first and more classic version is the Euclidean TSP (also known as the random point TSP), where the N cities are placed randomly and independently in d-dimensional Euclidean space, and the distances between cities are given by the Euclidean metric. The second is the random link TSP, largely developed within the context of disordered systems. In the random link TSP, it is not the positions of the cities but rather the lengths separating them that are independent and identically distributed random variables. We speak of "lengths" because there is no physical "distance" here, nor geometry of any sort, apart from a symmetry requiring that the length l_{ij} be equal to l_{ji} between any two cities i and j.[3] The great advantage of the random link TSP over the Euclidean (random point) TSP is that the lack of correlations among the lengths l_{ij} brings an analytical solution within reach.

It was noted by Vannimenus & Mézard (1984) that in the limit of large N, one may perfectly well choose the distribution of the individual l_{ij}'s in the random link TSP so as to match that of the d-dimensional Euclidean TSP. The probability of finding a point at Euclidean distance l is simply the surface area of a d-dimensional sphere of radius l:

$$\rho_d(l) = \frac{d\,\pi^{d/2}}{\Gamma(d/2+1)}\, l^{d-1}. \tag{I.3}$$

In this way, the random link and Euclidean models can be made to match one another, up to correlations between lengths. We may therefore consider the random link TSP as a *random link*

[2] Alternatively, the final constraint can be expressed by considering a partition of the N points into two sets A and B, and requiring that $\forall A, B,\ \sum_{i \in A, j \in B} n_{ij} \geq 1$.

[3] Asymmetric versions of the problem do exist as well, though we do not consider them here.

approximation to the Euclidean TSP. This approximation is far from perfect, since in reality correlations between Euclidean distances, such as the triangle inequality, are important. In terms of the results obtained, however, the random link approximation turns out to be a surprisingly good one.

The approach in this thesis is, first of all, to examine the d-dimensional Euclidean TSP, focusing on the large N behavior of the optimal tour length L_E ("E" for Euclidean). Our study is numerical, making use of good approximate methods for estimating L_E to within small, well-controlled errors. This, along with some theoretical insight into the scaling behavior to expect in arbitrary dimension d, enables us to obtain (in units where volume equals one) the large N *finite size scaling law* for the ensemble average of L_E:

$$\langle L_E(N,d)\rangle = \beta_E(d)\, N^{1-1/d}\left[1+\frac{A(d)}{N}+O\left(\frac{1}{N^2}\right)\right]. \tag{I.4}$$

At $d=2$ and $d=3$ we find extremely precise numerical values for $\beta_E(d)$: our results, $\beta_E(2) = 0.7120 \pm 0.0002$ and $\beta_E(3) = 0.6979 \pm 0.0002$, represent a precision 20 times greater than was possible using previously available methods.[4] It would of course be nice to have an analytical estimate of $\beta_E(d)$, and this is where the random link approximation comes in. As we have mentioned, this approximation is what makes the TSP analytically tractable. Using the "cavity" approach of Mézard, Parisi and Krauth (Mézard & Parisi, 1986*b*; Krauth & Mézard, 1989; Krauth, 1989), we obtain an estimate for $\beta_E(d)$. For $d=2$ and $d=3$, the random link approximation turns out to be a remarkably good one, with an error of less than 2%. In the $d\to\infty$ limit, we argue that in fact the approximation is exact up to $O(1/d^2)$. This leads to a conjecture on the behavior of $\beta_E(d)$ at large d, in terms of a series in $1/d$ where leading and subleading coefficients are given.

It is worth explaining why these Euclidean results are of broader interest than just to theoretical physicists. One of the earliest mathematical results for the Euclidean stochastic TSP (Beardwood, Halton & Hammersley, 1959) states that the random variable L_E obeys a self-averaging property:

$$\frac{L_E(N,d)}{N^{1-1/d}} \to \beta_E(d) \text{ as } N\to\infty \tag{I.5}$$

with probability 1. $\beta_E(d)$ is a well-defined quantity and not a random variable. This means that if we take a sequence of instances drawn randomly at large N, the relative fluctuations in the optimal tour length will be small from one instance to the next. (See Appendix B for an outline of the self-averaging proof.) Using this property, we may consider some large real-life Euclidean situation, such as an airplane serving numerous cities or a delivery truck serving a large number of retailers, and make the assumption that the instance "resembles" one from the random ensemble, *i.e.*, sites positioned independently. Knowing $\beta_E(d)$ allows us at least a rough estimate of how far any prospective tour in this instance is from "theoretical optimality". (A concrete example of this is given in Appendix C.)

[4] These values have subsequently been confirmed by Johnson, McGeoch & Rothberg (1996), using a modified version of our methods and more powerful computational resources. See Appendix E for further discussion.

From a theoretical point of view, the random link TSP is very much of interest as well, as a model of its own. The cavity method gives an analytical prediction for $\beta_{RL}(d)$, the random link counterpart of $\beta_E(d)$. However, this prediction is based on some rather subtle assumptions, reflecting work done in spin glasses. The central such assumption is that of *replica symmetry*, an ergodicity hypothesis stating that there is a unique thermodynamic limit, in a certain spin model onto which one maps the random link TSP. This assumption is known to be false in the case of spin glasses. For the TSP, however, numerical analyses performed in the $d = 1$ case (SOURLAS, 1986; KRAUTH & MÉZARD, 1989) suggest that it is in fact correct. We extend this numerical work to higher dimensions; our results indicate strongly that the cavity method gives correct predictions for all d.

Finally, there is a related stochastic problem, discussed in this thesis, that does not directly involve the TSP though it is motivated by the finite size scaling analysis used for the Euclidean TSP. This is a geometric issue concerning properties of distances between randomly distributed sites, notably distances from a point to its kth-nearest site. If the sites are distributed uniformly in Euclidean space (as cities are in our formulation of the Euclidean TSP), these kth-neighbor distances display a remarkable universality: they all have the same large N scaling law. Letting $\langle D_k(N)\rangle$ be the ensemble average of distances between kth-nearest neighbors, the N-dependence and the k-dependence of $\langle D_k(N)\rangle$ separate, and

$$\langle D_k(N)\rangle = \frac{[\Gamma(d/2+1)]^{1/d}}{\sqrt{\pi}} \frac{\Gamma(k+1/d)}{\Gamma(k)} \frac{\Gamma(N)}{\Gamma(N+1/d)} \tag{I.6}$$

in d-dimensional Euclidean space. If we work not in Euclidean space but on an arbitrary geometric manifold, a different universality emerges: when $\langle D_k(N)\rangle$ is written as a power of N times a correction series in $1/N$, the $O(1/N)$ term of the series is a topological invariant and does not depend on the detailed properties of the manifold. We discuss these universalities and surrounding issues, noting that some open questions include the application of this work to more complicated geometric problems such as triangulations.

• • •

In both its stochastic and non-stochastic forms, the TSP has attracted a large amount of research. In order to provide the reader with a better perspective of the context in which our own work is situated, let us survey some of the most relevant results in the literature.

The utopian ideal for a TSP algorithmic procedure would be one which finds the exact optimum, for an arbitrary instance, in a short amount of time. The fact that the problem is NP-hard appears to make this goal unattainable (see Appendix A for a discussion of the TSP's complexity). Nevertheless, computer scientist have for years tried to develop algorithms that reduce the amount of time necessary to find a guaranteed optimal tour. The most successful such algorithms have been based on the method of "branch-and-bound". The idea is as follows: a recursive procedure is used to break down an N-city problem into smaller subproblems. Each subproblem is then "relaxed" so as to find a bound. These bounds are subsequently used to eliminate further subproblems from consideration. In this way, the algorithm successively "branches" — divides problems into subproblems — and "bounds" — prunes branches that

violate bound restrictions (Balas & Toth, 1984). Finding efficient branching and relaxation procedures are, of course, highly sophisticated problems in themselves. A variant on this method known as "branch-and-cut" represents today's state of the art in exact algorithms: problems of up to 100 cities can typically be solved in minutes on a workstation (Padberg & Rinaldi, 1987).

However, when instances of larger sizes must be optimized, or a large sequence of instances must be optimized, exact algorithms are no longer practical. The next best choice, then, is to use approximate methods that find near-optimal tours quickly (typically in $O(N^2)$ operations). These are known as *heuristic* algorithms. For the tsp, the most powerful heuristics have been those that start with a non-optimal tour and successively attempt to improve upon that tour. The procedure generally used is known as *local search*: a certain number of links are removed from a tour and are reconnected so as to form a new and better tour; when all possible local changes have been tried without success, the resulting tour is then a local optimum. The simplest algorithm of this sort is 2-opt, where two links are deleted and the paths are reconnected differently. 3-opt does the same but using 3 links in each elementary move (Lin, 1965). Both heuristics tend to give tour lengths that are a few percent higher than the optimum for Euclidean tsp instances (Johnson & McGeoch, 1997). Various methods have been popular for improving on this. The simplest is to repeat 2-opt or 3-opt from different random starting tours, taking the best local optimum found over many trials, so as to cover more of the energy landscape (see Figure I–2). Another method is simulated annealing, which can outperform repeated 3-opt (in terms of both running time and percent excess of optimum) only for sufficiently large instance sizes ($N > 1000$). The most successful of the local search heuristics, however, is that of Lin & Kernighan (1973). This algorithm uses a "variable-depth neighborhood search" method where a large number of links can be changed at once, but where redundancy in searching is reduced to the point where a million-city Euclidean instance can be solved in less than an hour, with less than 2% excess over the optimum (see Appendix D for a description of the heuristic; see also Johnson & McGeoch, 1997).

Improving upon Lin-Kernighan has been a goal in algorithmic research for over 20 years, with relatively limited success. One particularly creative attempt has been through the use of *genetic algorithms* where, inspired by the process of evolution, one attemps to establish new starting tours by "mating" (combining aspects of two "parent" locally optimal tours) and by "mutation" (randomly altering a locally optimal tour). Although the mating principle has not yet proven to produce a competitive heuristic, the mutation principle has been used with success in the "metaheuristic" known as Chained Local Optimization (clo) by Martin, Otto & Felten (1991). clo involves successively applying a random "kick" to a locally optimal tour, and then using Lin-Kernighan to locally optimize from there. If the new tour is better than the old one, the procedure iterates from that point; if not, the old tour is again used and a new random "kick" is attempted. (A qualitative description may be found in Appendix D.) clo outperforms Lin-Kernighan for instances of $N > 100$, taking seconds to get within 1% of the optimum for a 1000-city instance as opposed to minutes for the equivalent performance simply repeating Lin-Kernighan from random starts (Martin & Otto, 1996).

All of what has been mentioned so far applies to optimizing individual tsp instances. When we speak of the stochastic tsp, we must access an entire ensemble of instances. One approach

is simply to run optimization heuristics over a large sample from the ensemble, thus obtaining results to within a quantifiable statistical error. This approach is indeed the basis of most numerical studies, including our own. These numerical studies, however, have also been accompanied by a certain amount of theoretical work.

For the Euclidean stochastic TSP, we have already seen the *self-averaging* property shown by Beardwood, *et al.* These authors, in the same article, gave lower and upper bounds on the asymptotic limit $\beta_E(d)$. The lower bound is simple: it uses the fact that, at best, any city on the tour has links to its two nearest neighbors. Nearest neighbor distances may be calculated over the ensemble (as we discuss in depth in Chapter IV), and one finds that $\beta_E(2) \geq 5/8$. It is straightforward to generalize this bound to any dimension d. For the upper bound, a method known as "strip" was used, whereby the space is divided into columns (strips) and a suboptimal tour is constructed by visiting cities alternately up and down the columns. In 2 dimensions, this suboptimal tour length may be calculated exactly over the ensemble; the result is a complicated expression involving special functions, but may be numerically computed as $\beta_E(2) \leq 0.9204$.

No better exact bounds are known today; the best that has been done is to use heuristic methods on very large instances (or many very large instances) to establish estimated bounds. In this way, an approximation to the lower bound of HELD & KARP (1970), involving a linear programming relaxation of the TSP, has very recently been used by JOHNSON, MCGEOCH & ROTHBERG (1996) to estimate $\beta_E(2) \geq 0.708$. LEE & CHOI (1994) have used a variant on simulated annealing to obtain the upper bound $\beta_E(2) \leq 0.721$. At $d > 2$, virtually no improvements over the exact bounds have existed prior to our own work. (Johnson *et al.* have, however, in their work on the Held-Karp bound, successfully applied a modified version of our numerical methods to higher precision, confirming our own $d = 2$ and $d = 3$ results and extending to the $d = 4$ case.) The only other related large d result is an unproven conjecture by BERTSIMAS & VAN RYZIN (1990), claiming that when $d \to \infty$, $\beta_E(d) \sim \sqrt{d/2\pi e}$.

For the random link stochastic TSP, results have been largely due to the work of theoretical physicists. VANNIMENUS & MÉZARD (1984) and KIRKPATRICK & TOULOUSE (1985) first considered the statistical mechanics of the TSP, using analytical tools stemming from the study of spin glasses. Vannimenus and Mezard showed that the equivalent of the Bertsimas-van Ryzin conjecture is correct for the random link case, *i.e.*, $\beta_{RL}(d) \sim \sqrt{d/2\pi e}$ when $d \to \infty$. Kirkpatrick and Toulouse studied several low-temperature macroscopic quantities of the random link model, analyzing the number of links in common between local optima at finite temperature. This "overlap" leads to the notion of a state-space *bond distance*, measuring by how much two tours differ from each other. In analogous spin glass problems, such distances between local minima display a property known as *ultrametricity*, a stronger form of the triangle inequality, requiring that for any three distances at least two must be equal (they form an isoceles triangle) and the third must be less than or equal to the other two. Non-trivial ultrametricity is considered a typical sign of a lack of ergodicity in spin glasses, called *replica symmetry breaking*. SOURLAS (1986), numerically extending Kirkpatrick and Toulouse's low-temperature analysis, found on the basis of the tour overlap that ultrametricity does not play a role in the $d = 1$ random link TSP, and that replica symmetry most likely holds. In the process, he observed that solutions only contain links between very near neighbors, and used this property to motivate an improved

optimization heuristic that works by disallowing certain links.

Under the assumption that the replica symmetric solution is the correct one, MÉZARD & PARISI (1986*b*) and KRAUTH & MÉZARD (1989) used the cavity method to find an analytical expression for $\beta_{RL}(d)$, in the form of an integral equation. They performed further numerical checks at $d = 1$ (through direct simulations of the random link TSP) that confirmed their analytical cavity results, and briefly discussed the $d = 2$ result. We are aware of no further attempts, prior to our own, to extend the analysis of the cavity solutions to higher dimensions.

• • •

Let us conclude this chapter with a brief guide to the layout of what follows in the thesis. Chapter II deals with scaling laws in the stochastic TSP. In the articles presented in the chapter we discuss our most important results, notably a numerical study of the finite size scaling of L_E, and an analytical study of the dimensional scaling of $\beta_E(d)$ by way of the random link approximation. Our analysis of $L_E(N, d)$ at large N is, to our knowledge, the first to use a systematic method for extrapolating to the large N limit. We are thus able to improve significantly on previous estimates of $\beta_E(d)$. Using the random link approximation, we also present evidence supporting the Bertsimas-van Ryzin conjecture that $\beta_E(d) \sim \sqrt{d/2\pi e}$ at large d, and propose a much stronger conjecture: that

$$\beta_E(d) = \sqrt{\frac{d}{2\pi e}}\,(\pi d)^{1/2d}\left[1 + \frac{2 - \ln 2 - 2\gamma}{d} + O\left(\frac{1}{d^2}\right)\right], \tag{I.7}$$

where γ is Euler's constant ($\gamma \approx 0.57722$).

Chapter III covers the random link TSP. We discuss the background of the analytical tools used, and the physical model serving as the basis of the cavity method. Through simulations, we then test the cavity results for $d > 1$, and find good numerical justification for the hypothesis of replica symmetry.

Chapter IV provides a discussion of the universality in kth-nearest neighbor distances for randomly placed sites in Euclidean space, namely that the same scaling in N applies to the mean distance $\langle D_k(N)\rangle$ regardless of k. We generalize this study to arbitrary geometric manifolds, and find that, although the scaling of $\langle D_k(N)\rangle$ does depend on k in general, its $O(1/N)$ term is a topological invariant and does not depend on the precise shape of the manifold. We note that these properties, though arising from relatively simple physical arguments, can be applied to far more complex geometric problems.

In each of these subsequent three chapters, an overview of the subject is given, following which one or more articles make up the body of the chapter. In Chapter II, these are articles that have already been published, and are presented in reprint form. In Chapters III and IV, the articles have not yet been submitted, and so it has been possible to take some liberties with their presentation, following wherever possible the style adopted in the rest of the text. Not being constrained by editorial restrictions concerning brevity, we have also attempted to adopt a more pedagogical tone than will ultimately be permitted in published form.

Several appendices follow the main body of the thesis. In Appendix A, we define the notions P and NP as used in computational complexity theory. We show how to phrase the TSP in such

a way that it is an NP-hard problem, and discuss polynomial time approximation algorithms. Appendix B covers self-averaging in $L_E(N,d)$; we outline a simple proof of this property, establishing that $L_E(N,d)/N^{1-1/d}$ goes to a constant $\beta_E(d)$ when $N \to \infty$. In Appendix C we explain how our Euclidean stochastic TSP results, although intended for a uniform distribution of cities, may easily be applied to non-uniform cases. We give the example of the AT&T-532 instance (containing 532 AT&T telecommunications sites around the U.S.), and show that we come within 1% of the true optimum by regarding it as a member of the stochastic ensemble. In Appendix D, we discuss our numerical methodology, giving an explanation of the workings of the Lin-Kernighan and CLO algorithms for optimizing TSP instances. Finally, in Appendix E, we describe a more recent numerical study of $\beta_E(d)$ and $\beta_{RL}(d)$ performed by JOHNSON, MCGEOCH & ROTHBERG (1996), using a variant on our methods. They obtain results in close agreement with ours, and by combining their data and ours we are in fact able to refine our $\beta_E(2)$ estimate slightly further.

Chapter II

Scaling laws in the Euclidean TSP

Contents

Chapter overview 17
Reprint: Physical Review Letters 21
Reprint: Journal de Physique 27

– *Chapter overview* –

Historically, the traveling salesman problem has been of greatest interest in its Euclidean formulation. In this form of the problem, each of the N cities has a position in d-dimensional space (in practice, most often 2-dimensional space), and the lengths separating pairs of cities are defined to be the associated Euclidean distances. The traveling salesman tour is thus a tour through a metric space.

Much past work on this problem has concentrated on developing algorithms to solve certain large instances, or often certain classes of instances. One goal in operations research has been to find methods of reducing redundancy in exhaustive search — methods that, while still exponentially slow in N, at least have a small enough coefficient in the exponential that they find the exact optimum in a reasonable amount of time. Another goal has been to classify special cases of instances. The idea here is to identify properties of the layout of cities that could guarantee finding an optimal tour within, say $O(N^\alpha)$ steps for some fixed α. Finally, considerable efforts have been devoted to finding mathematically rigorous bounds on tour lengths.

In the case of the stochastic Euclidean TSP, however, the emphasis is somewhat different. The positions of the cities in the tour are now independent random variables; in the model that we consider here, the distribution is uniform over space.[1] It is no longer a matter of optimizing a single instance. We are now working with an entire ensemble of instances, and the problem is to find properties of the optimal tour length L_E (the actual path that tours follow is of less interest when the positions of the cities are not fixed quantities) considered in the ensemble of all possible instances. The challenge consists of understanding what happens at large instance sizes. The fundamental property here is that of *self-averaging*, proven by Beardwood, Halton & Hammersley (1959): as N gets large the relative fluctuations in the optimal tour length go to zero, and the distribution of L_E, up to a scaling factor, becomes more and more sharply peaked.

[1] The non-uniform case is discussed in Appendix C.

The real quantity of interest is then the *mean* optimal tour length over the ensemble — at a given instance size N and for a given dimension d — which we denote $\langle L_E(N,d)\rangle$, using units where the volume of our space is 1.

An outline of a more recent and simpler proof of self-averaging, due to KARP & STEELE (1985), is given in Appendix B. The precise statement of the property is that with probability 1, for any sequence of random instances, $L_E(N,d)/N^{1-1/d}$ will approach the instance-independent quantity $\beta_E(d)$ at large N. Note that the scaling factor $N^{1-1/d}$ may also be seen by a simple physical argument: we are working in a fixed volume, so the mean volume *per city* scales as N^{-1}, and so the mean distance between cities scales as $N^{-1/d}$; the tour contains N such links, hence its length scales as $N^{1-1/d}$.

The question then becomes how to determine the limiting

$$\beta_E(d) \equiv \lim_{N\to\infty} \frac{\langle L_E(N,d)\rangle}{N^{1-1/d}}. \tag{II.1}$$

For small values of d, this may be answered by combining numerical methods with some theoretical insights — namely, *how* $L_E(N,d)/N^{1-1/d}$ converges to its large N limit. We thus consider the finite size scaling of $L_E(N,d)$. At large d, this no longer feasible numerically. Fortunately, in the limit $d\to\infty$, an analytical approach may be developed for obtaining $\beta_E(d)$ to within an excellent approximation.

What sort of numerical methods are needed for a finite size scaling analysis? At a given d, we must find $\langle L_E(N,d)\rangle$ over many values of N. At any particular value of N, in order to obtain this quantity we must, in turn, average $L_E(N,d)$ over a large number of instances. (The smaller N is, the more instances we will have to consider in order to minimize the effects of fluctuations.) At any given instance, finally, in order to find the optimum we must have efficient algorithms. Unfortunately, exact algorithms are too slow to be of use here, requiring minutes of CPU time to solve typical instances of $N = 100$; optimizing a statistically significant sample of random instances in this way could take weeks of computing time.

There is, fortunately, a better approach. Powerful *heuristic* algorithms exist — inexact algorithms which find "good" tours but not always the optimal one. The most suitable method is that of *local search*, where the algorithm takes a non-optimal tour and iteratively improves upon it so as to bring it closer to optimality. We may think of this in terms of the landscape in Figure I–2; the algorithm finds the local minimum of the valley in which we start, and if it is a good algorithm, the landscape will be such that there are a small number of broad valleys. As we execute a local search heuristic increasingly many times on a given instance, each time starting from a different random initial tour, we then explore increasing parts of the landscape and find the true optimum with increasing probability. Moreover, even with multiple runs, these heuristics work considerably faster than exact algorithms — especially at large N where running time grows only polynomially and not exponentially in N. This means that in a given amount of time, we are able to optimize a much larger sample of instances (and thus have a far smaller statistical error) using heuristics than using exact algorithms. The net effect is counterintuitive: because of computing time considerations, we actually get *more* precise results using these inexact algorithms than using exact algorithms. There is, admittedly, a systematic

bias in the results arising from the use of quasi-optimal tour lengths: no matter how many times we run a heuristic, we are never *sure* to have found the global optimum in the end. However, we choose the number of runs per instance sufficiently conservatively that this systematic bias — estimated for each value of N on the basis of a number of test instances — is so small as to be negligible compared with the statistical error. (Our numerical methodology is discussed in Appendix D.)

Equipped with this numerical approach and certain theoretical notions concerning the scaling to expect in N — described in detail in the articles that follow — we find the large N finite size scaling law:

$$\langle L_E(N,d)\rangle = \beta_E(d)\, N^{1-1/d}\, \left[1 + \frac{A(d)}{N} + O\left(\frac{1}{N^2}\right)\right] \tag{II.2}$$

using a d-dimensional unit hypercube with periodic boundary conditions.[2] Understanding the scaling in this manner enables us to obtain $\beta_E(d)$ to high precision at $d = 2$ and $d = 3$ (the cases that we consider numerically): $\beta_E(2) = 0.7120 \pm 0.0002$ and $\beta_E(3) = 0.6979 \pm 0.0002$. When d gets larger, however, the feasibility of numerical solutions rapidly decreases. In order to get an idea of the dimensional scaling of $\beta_E(d)$, we therefore turn to the analytical approach used in the *random link approximation*.

The random link approximation consists of assuming that lengths l_{ij} between city i and city j ($i < j$) are completely independent of one another. Thus under the random link approximation (which could perhaps more properly be called the "independent link approximation") correlations between distances, such as the triangle inequality, are neglected. Making this approximation allows us to take advantage of an analytical approach developed by Mézard & Parisi (1986*b*), Krauth & Mézard (1989) and Krauth (1989). We use this approach to obtain $\beta_{RL}(d)$, the random link value approximating $\beta_E(d)$. $\beta_{RL}(d)$ arises from a system of integral equations and does not appear to have an exact analytical solution, but can on the other hand be solved numerically to arbitrary precision. For small values of d, we find that $\beta_{RL}(d)$ is a good approximation to $\beta_E(d)$: for $d \leq 3$, the relative error is within about 2%.

For large d we are in fact able to express $\beta_{RL}(d)$ analytically, via a power series in $1/d$. To leading and subleading order, this gives

$$\beta_{RL}(d) = \sqrt{\frac{d}{2\pi e}}\, (\pi d)^{1/2d} \left[1 + \frac{2 - \ln 2 - 2\gamma}{d} + O\left(\frac{1}{d^2}\right)\right], \tag{II.3}$$

where γ represents Euler's constant ($\gamma \approx 0.57722$). We then present an argument, based on a theoretical analysis of how certain heuristic algorithms operate, suggesting that the relative error in the random link approximation itself decreases at least as fast as $O(1/d^2)$ in the limit $d \to \infty$. This enables us to conjecture, finally, that the large d expression (II.3) applies to $\beta_E(d)$ as well, and thus gives the correct dimensional scaling for the Euclidean optimal tour length.

[2] Jaillet (1993) has proven that $\beta_E(d)$ is in fact the same regardless of whether periodic or open boundary conditions are used.

Reprint: Physical Review Letters

VOLUME 76, NUMBER 8 PHYSICAL REVIEW LETTERS 19 FEBRUARY 1996

Finite Size and Dimensional Dependence in the Euclidean Traveling Salesman Problem

Allon G. Percus* and Olivier C. Martin†

Division de Physique Théorique, Institut de Physique Nucléaire, Université Paris-Sud, F-91406 Orsay Cedex, France

(Received 20 July 1995)

We consider the Euclidean traveling salesman problem for N cities randomly distributed in the unit d-dimensional hypercube, and investigate the finite size scaling of the mean optimal tour length L_E. With toroidal boundary conditions we find, motivated by a remarkable universality in the kth nearest neighbor distribution, that $L_E(d=2) = (0.7120 \pm 0.0002)N^{1/2}[1 + O(1/N)]$ and $L_E(d=3) = (0.6979 \pm 0.0002)N^{2/3}[1 + O(1/N)]$. We then consider a mean-field approach in the limit $N \to \infty$ which we find to be a good approximation (the error being less than 2.1% at $d = 1, 2$, and 3), and which suggests that $L_E(d) = N^{1-1/d}\sqrt{d/2\pi e}\ (\pi d)^{1/2d}[1 + O(1/d)]$ at large d.

PACS numbers: 02.60.Pn, 02.70.Lq, 64.60.Cn

The traveling salesman problem (TSP) is one of the best known combinatorial optimization problems. It is *NP* complete (suggesting that no algorithm exists for solving the problem in polynomial time), and it serves as a fertile ground for analytical and numerical approaches to optimization problems in general. It is also one of the few optimization problems that have been studied extensively in the context of statistical mechanics.

The TSP, as we consider it, is as follows: Given N points ("cities") in a space, the problem is to find the length of the shortest closed path ("tour") going through each city exactly once. Two particular forms of the problem have been investigated in depth. The first, which has attracted the most attention among computer scientists and mathematicians, is the Euclidean TSP: The N cities are randomly distributed in a d-dimensional hypercube and the distances between cities are given by the Euclidean metric. The second, which has been of particular interest within the statistical physics community, is the random link TSP: The lengths l_{ij} separating cities i and j are taken as independent random variables with a given distribution $\rho(l)$.

It has been noted by Mézard and Parisi [1] that the random link model, with $\rho(l)$ appropriately chosen, maps onto the Euclidean model if correlations between three or more distances are neglected (no triangle inequality, for instance). This suggests that the random link TSP can be considered as a mean-field approximation to the Euclidean case, and perhaps that this approximation becomes exact in the limit $d \to \infty$.

Our intention in this Letter is twofold. First, for the Euclidean TSP we investigate finite size corrections to the mean optimal tour length L_E, in the large N ("thermodynamic") limit. To our knowledge there has been no prior work on this subject, in spite of a great deal of interest in L_E in the thermodynamic limit itself. Second, we explore the dimensional dependence of L_E using a mean-field approach (the random link TSP in conjunction with the "cavity method" [1,2]). We extend the work of Krauth and Mézard [3] to find the mean-field optimum L_{MF} in the thermodynamic limit, as a function of dimension. Comparing mean-field results with Euclidean $N \to \infty$ results at low d shows that mean field does considerably better than previously expected, and suggests that in quite natural units, L_E can be written as a power series in $1/d$.

Euclidean model: Finite size scaling ($d = 2$).—We start with the case of N cities distributed randomly and uniformly in a unit square. Numerous heuristic approaches have been developed to find near-optimal TSP tours given a particular configuration ("instance") of cities. For our purposes, the most convenient methods are local-optimization heuristics such as the Lin-Kernighan (LK) [4] and the chained local optimization (CLO) [5] algorithms. With these algorithms, repeated runs on a given instance using different random starts produce the optimal tour with increasing probability.

It has been shown [6] that in the large N limit the optimal tour length *for a given instance* $\tilde{L}_E$ is self-averaging up to a scaling factor

$$\lim_{N\to\infty} \frac{\tilde{L}_E}{N^{1-1/d}} = \beta_E\,,$$

where convergence to the instance-independent β_E is with probability 1 (in the ensemble of instances with randomly distributed cities). Much past work has concentrated on optimizing single instances at large N (see [5,7,8]). Here, however, our concern is to calculate β_E along with an estimate of statistical error, and so instead we average over a large number of instance. There is necessarily a tradeoff in the choice of N: At small N alone we cannot confidently predict the finite size scaling behavior, whereas at large N the large amount of computing time necessary for each optimization sharply limits the number of instances we can optimize reliably, and increases the statistical error. We therefore choose several small values of N ($N = 12$ through $N = 17$) where we optimize using LK, and two larger values ($N = 30$ and $N = 100$) where we optimize using CLO.

Given $L_E(N)$ at different values of N, then, we wish to extrapolate and extract the limit β_E, as well as finite size corrections. In order to eliminate the effects of

 0031-9007/96/76(8)/1188(4)$06.00

surface terms, we use periodic boundary conditions in the Euclidean distance metric. An indication of the size dependence to be expected in $L_E(N)$ may be found by looking at the distance D_k between kth nearest neighbors, averaged over the ensemble of instances. A direct calculations shows that, given N cities distributed randomly and uniformly over the d-dimensional unit hypercube (with periodic boundary conditions),

$$D_k(N,d) \sim \binom{N-1}{k-1}(N-k)d\left[\frac{\pi^{d/2}}{\Gamma(d/2+1)}\right]^k \times \int_0^{1/2} r^{dk}\left[1-\frac{\pi^{d/2}}{\Gamma(d/2+1)}\right]^{N-k-1} dr\,,$$

where exponentially small corrections in N have been neglected.

Recognizing this integral (up to a change of variable and further exponentially small corrections in N) as a beta function, we find that

$$D_k(N,d) \sim \frac{\Gamma(N)}{\Gamma(N+1/d)}\,\frac{\Gamma(d/2+1)^{1/d}}{\sqrt{\pi}}\,\frac{\Gamma(k+1/d)}{\Gamma(k)}\,. \tag{1}$$

Notice that there is a complete separation here of the N dependence and the k dependence. This is indeed a surprising universality: It means that up to exponentially small corrections, *all kth nearest neighbor mean distances have exactly the same scaling law in N,* namely, $\Gamma(N)/\Gamma(N+1/d)$. It might be expected, then, that the length of a TSP tour consisting of N links would have large N scaling behavior

$$N\frac{\Gamma(N)}{\Gamma(N+1/d)} = N^{1-1/d} \times\left[1+\frac{1/d-1/d^2}{2N}+O\left(\frac{1}{N^2}\right)\right],$$

where the right-hand side follows from Stirling's formula.

In fact, due to correlations between k and N in the optimal tour, this is not quite the case. Figure 1 shows our results for L_E divided by the scaling quantity above, at $d = 2$: We find that this is, to a good fit, itself a power series in $1/N$, albeit one with a small first-order term. The asymptotic $N \to \infty$ value is $\beta_E = 0.7120 \pm 0.0002$, where the error is obtained on the basis of χ^2 analysis. This result is, to our knowledge, the most precise to date for the Euclidean TSP in the thermodynamic limit.

The methods by which we obtained the results in Fig. 1 are themselves of some importance. For runs optimized by LK ($N = 12$ through $N = 17$), we averaged over the results of 250 000 instances, where for each instance we took the best (lowest) optimum found in ten random starts (ten different runs). For $N = 30$ we averaged over 10 000 instances, taking for each one the best optimum found by CLO (ten Monte Carlo iterations per run) in five random starts. For $N = 100$ we averaged over 6000 instances, taking for each one the best optimum found by CLO (ten Monte Carlo iterations per run) in twenty random starts. These methods introduce a systematic error, because they do not always find the true optimum; we estimated this error by performing a large number of runs on a few instances and measuring the average expected error (weighted by the probability of making that error when choosing the best out of ten random starts). In all cases, we verified that the systematic error stayed under 10% of the statistical error shown in the error bars.

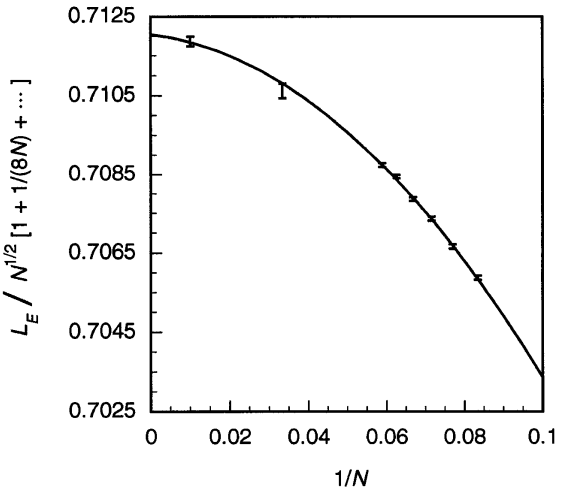

FIG. 1. Finite size dependence of rescaled Euclidean 2D TSP optimum. Best fit ($\chi^2 = 5.48$) is given by $L_E/N^{1/2}[1 + 1/(8N) + \cdots] = 0.7120(1 - 0.0171/N - 1.048/N^2)$. Error bars represent statistical errors.

In order to reduce the statistical noise further, we used the following variance reduction method: Recognizing that $L_B(N) \equiv N(D_1 + D_2)/2$ is a lower bound on the tour length (each city is at best connected to its first- and second-nearest neighbors), write the estimator for L_E as $\langle \tilde{L}_E - \lambda \tilde{L}_B\rangle + \lambda L_B$. $\tilde{L}_E$ and $\tilde{L}_B$ denotes values *for a particular instance,* the angular brackets represent the average over instances sample, and the ensemble average L_B can be calculated analytically [see Eq. (1)]. λ is a parameter which we adjust to minimize the variance of our new estimator. In practice, optimal values of λ ($\lambda \approx 0.75$) enabled us to reduce the error by over 60%. Other variance reduction methods can also be used [9], but ours has the advantage of introducing no new systematic error.

Mean-field method.—We now turn our attention to the mean-field approximation, based on the random link TSP. Rather than having N cities distributed randomly in a hypercube, we now have lengths l_{ij} between cities i and j ($1 \le i < j \le N$) distributed as *independent* random variables according to a certain distribution $\rho(l)$. We take $\rho(l)$ to be the probability distribution of lengths between cities in the d-dimensional Euclidean problem, in the absence of finite size effects:

$$\rho(l) = d\pi^{d/2} l^{d-1}/\Gamma(d/2+1)\,.$$

This establishes a mapping in the thermodynamic limit between the random link TSP and the Euclidean TSP, neglecting all correlations among (Euclidean) distances.

The mean-field "model" is the random link TSP, described for our purposes by the "cavity equations" written down by Krauth and Mézard [3]. In our language this leads to

$$\beta_{\mathrm{MF}}(d) = \frac{1}{\sqrt{\pi}}\,\frac{d}{2}\left[\frac{\Gamma(d/2+1)}{\Gamma(d+1)}\right]^{1/d} \times \int_{-\infty}^{\infty} G_{d-1}(x)\,[1+G_{d-1}(x)]\,e^{-G_{d-1}(x)}\,dx\,,$$

where $\beta_{\mathrm{MF}} \sim L_{\mathrm{MF}}/N^{1-1/d}$ as in the Euclidean case, and

$$G_d(x) = \int_{-x}^{\infty} \frac{(x+y)^d}{d!}\,[1+G_d(y)]\,e^{-G_d(y)}\,dy\,. \quad (2)$$

It has been argued persuasively, notably on the basis of excellent agreement in the $d = 1$ case [3], that the cavity method is exact for the $N \to \infty$ random link TSP. In the following discussion we shall also present further justification for this assumption.

There is no known analytical solution of the integral equation for $G_d(x)$ given in Eq. (2). However, it can be solved numerically; this was done by Krauth and Mézard at $d = 1$ and $d = 2$, giving $\beta_{\mathrm{MF}}(d=1) = 1.0208$ and $\beta_{\mathrm{MF}}(d=2) = 0.7251$ [3]. These values may be compared with $\beta_E(d)$: Under periodic boundary conditions $\beta_E(d=1) = 1$ (trivially) and $\beta_E(d=2) = 0.7120$ (see previous section). Therefore, at $d = 1$ mean field has a 2.1% excess with respect to the Euclidean value, and at $d = 2$ a 1.8% excess (see also Table I). Already at low dimension, then, mean field gives quite a good approximation to the Euclidean case. It is amusing to note that Krauth and Mézard themselves assumed a rather inaccurate Euclidean value $\beta_E(d=2) = 0.749$, and so their mean-field results seemed poorer to them than they actually were.

We now extend the numerical solution of Eq. (2) to higher dimensions. As in the problem of Euclidean finite size scaling, we can get an indication of what dimensional dependence to expect in $L_{\mathrm{MF}}(d)$ by looking at the mean kth nearest-neighbor distance D_k multiplied by the number of links N. In the thermodynamic limit, Eq. (1) gives

$$ND_k(d) \sim \begin{cases} N^{1-1/d}\dfrac{\Gamma(d/2+1)^{1/d}}{\sqrt{\pi}} \\ \quad \times \dfrac{\Gamma(k+1/d)}{\Gamma(k)}, \\ N^{1-1/d}\sqrt{\dfrac{d}{2\pi e}}\,(\pi d)^{1/2d} \\ \quad \times\left[1+O\left(\dfrac{\ln k}{d}\right)\right] \end{cases} \quad \text{at large } d\,.$$

Dividing by $N^{1-1/d}$, this suggests that

$$\beta(d) = \sqrt{\frac{d}{2\pi e}}\,(\pi d)^{1/2d}\left[1+O\left(\frac{1}{d}\right)\right].$$

TABLE I. Comparison of Euclidean and mean-field TSP optima (rescaled) at dimension up to $d = 3$.

d	β_E	β_{MF}	MF % excess
1	1	1.0208	+2.1%
2	0.7120 ± 0.0002	0.7251	+1.8%
3	0.6979 ± 0.0002	0.7100	+1.7%

Figure 2 shows that this is indeed so for the mean-field results obtained by numerical resolution of Eq. (2). Looking at $\beta_{\mathrm{MF}}/\sqrt{d/2\pi e}\,(\pi d)^{1/2d}$, we find an excellent fit by a $1/d$ power series with a leading order term which, to the precision of our raw numerical data, is indistinguishable from 1.

The fact that $\beta_{\mathrm{MF}}/\sqrt{d/2\pi e}$ at $d \to \infty$ is another confirmation of the validity of the cavity method, as this property is known to be true for the pure random link TSP [10]. We have thus added to Krauth and Mézard's investigation (at $d = 1$) further evidence (at $d \to \infty$) that the cavity method is exact.

Finally, let us rewrite the left-hand side of the best-fit equation in Fig. 2 with an additional $(1/2)^{1/2d}$ factor in the denominator:

$$\frac{\beta_{\mathrm{MF}}}{\sqrt{d/2\pi e}\,(\pi d/2)^{1/2d}} = 0.999997 + \frac{0.499395}{d} + O\left(\frac{1}{d^2}\right).$$

Notice that the $1/d$ coefficient is practically indistinguishable from $1/2$. An interpretation of this remarkable result is given in [11].

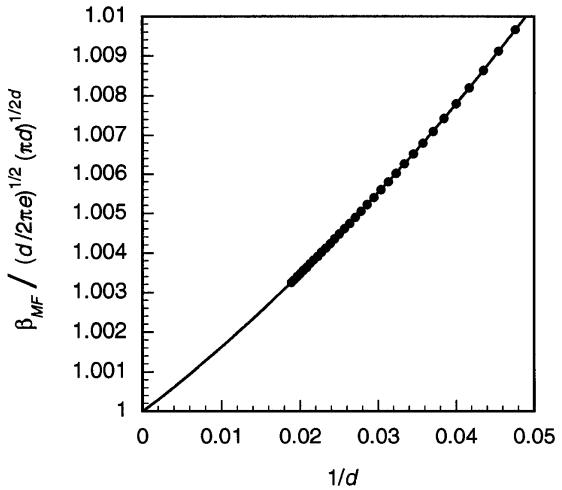

FIG. 2. Dimensional dependence of rescaled mean-field TSP optimum. Best fit ($\chi^2 = 7.46 \times 10^{-11}$) is given by $\beta_{\mathrm{MF}}/\sqrt{d/2\pi e}\,(\pi d)^{1/2d} = 0.999997 + 0.152821/d + 1.05488/d^2$.

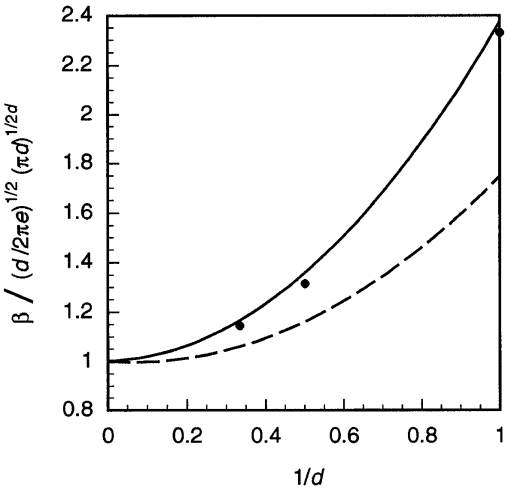

FIG. 3. Rescaled Euclidean TSP optimum (points) as a function of dimension, sandwiched between mean-field optimum (solid line) and exact lower bound (dashed line).

Euclidean model: Dimensional dependence.—Given the mean-field results, we now return to the Euclidean model. Table I shows the numerical result at $d = 3$ (obtained by the same heuristic methods as in the $d = 2$ case) together with the mean-field value, and the $d = 1$ and $d = 2$ results presented earlier.

These results suggest that β_{MF} is an upper bound for β_E (and heuristic arguments [11] also provide support for this). At the same time, there is the strict lower bound $L_B \equiv N(D_1 + D_2)/2$, mentioned earlier in the discussion on variance reduction. Figure 3 shows the Euclidean results "sandwiched" between the corresponding mean field and lower bound quantities, both of which may be written in the $d \to \infty$ limit as $\beta(d) = \sqrt{d/2\pi e}\,(\pi d)^{1/2d}[1 + O(1/d)]$. We conjecture that mean field does indeed remain an upper bound at all values of d, and consequently that β_E behaves asymptotically at large d as

$$\beta_E(d) = \sqrt{\frac{d}{2\pi e}}\,(\pi d)^{1/2d}\left[1 + O\left(\frac{1}{d}\right)\right].$$

This would also support a weaker conjecture by Bertsimas and van Ryzin [12], stating that for the Euclidean TSP, $\beta_E \sim \sqrt{d/2\pi e}$ at $d \to \infty$.

In conclusion, we have investigated the finite size behavior of the Euclidean TSP optimum under periodic boundary conditions, and have seen that at $d = 2$, L_E converges as a $1/N$ series:

$$\frac{L_E}{N^{1/2}[1 + 1/(8N) + \cdots]} = \beta_E\left(1 - \frac{0.0171}{N} - \cdots\right).$$

In the process we have extracted what we believe to be the best result to date for the thermodynamic limit: $\beta_E(d = 2) = 0.7120 \pm 0.0002$.

At the same time we have, by means of a mean-field method, examined the dimensional dependence of the TSP. We have found that mean field is a good approximation ($<$ 2.1% error) to the Euclidean TSP at $d = 1$, 2, and 3. We have seen numerically that at $d \to \infty$ the cavity equations are compatible with the exact random link TSP result, and thus have provided further evidence that they are exact at all dimensions. Additional work is in progress to understand the coefficient $1/2$ in the subleading term of the cavity equation solution. Finally, comparing our mean-field and Euclidean results suggests not only that the Bertsimas-van Ryzin conjecture for the large d limit of $\beta_E(d)$ is correct, but also that the asymptotic behavior is in fact $\beta_E(d) = \sqrt{d/2\pi e}\,(\pi d)^{1/2d}[1 + O(1/d)]$.

We are grateful to N. Cerf for his contributions, as well as to O. Bohigas for having suggested the present research. We also acknowledge fruitful discussions with E. Bogomolny, M. Mézard, S. Otto, and N. Sourlas. O. C. M. acknowledges support from NATO travel Grant No. CRG 920831. Division de Physique Théoriques is a unité de Recherche des Universités Paris XI et Paris VI associée au CNRS.

Note added.–Since our submission, D. S. Johnson *et al.* [13], using slightly different methods, have found values for $\beta_E(d)$ compatible with ours at $d = 2$ and 3.

*Electronic address: percus@ipncls.in2p3.fr
†Electronic address: martin_o@ipncls.in2p3.fr

[1] M. Mézard and G. Parisi, Europhys. Lett. **2**, 913 (1986).
[2] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
[3] W. Krauth and M. Mézard, Europhys. Lett. **8**, 213 (1989).
[4] S. Lin and B. Kernighan, Oper. Res. **21**, 498 (1973).
[5] O. Martin, S. W. Otto, and E. W. Felten, Oper. Res. Lett. **11**, 219 (1992).
[6] J. Beardwood, J. H. Halton, and J. M. Hammersley, Proc. Cambridge Philos. Soc. **55**, 299 (1959).
[7] D. S. Johnson, in *Proceedings of the 17th Colloquium on Automata, Language, and Programming* (Springer-Verlag, Berlin, 1990), p. 446.
[8] W. D. Smith, Ph.D. thesis, Princeton University, 1989.
[9] N. Sourlas, Europhys. Lett. **2**, 919 (1986).
[10] J. Vannimenus and M. Mézard, J. Phys. Lett. (Paris) **45**, 1145 (1984).
[11] N. J. Cerf, J. H. Boutet de Monvel, O. Bohigas, O. C. Martin, and A. G. Percus (to be published).
[12] D. J. Bersimas and G. van Ryzin, Oper. Res. Lett. **9**, 223 (1990).
[13] D. S. Johnson, L. A. McGeoch, and E. E. Rothberg, in Proceedings of the 7th Annual ACM-SIAM Symposium on Discrete Algorithms (to be published).

Reprint: Journal de Physique

J. Phys. I France **7** (1997) 117–136 JANUARY **1997**, PAGE 117

The Random Link Approximation for the Euclidean Traveling Salesman Problem

N.J. Cerf (*), J. Boutet de Monvel, O. Bohigas, O.C. Martin and A.G. Percus (**)

Division de Physique Théorique (***), Institut de Physique Nucléaire, Université Paris-Sud, 91406 Orsay Cedex, France

(Received 22 April 1996, revised 28 June 1996, accepted 23 September 1996)

PACS.02.60.Pn – Numerical optimization
PACS.02.70.Lq – Monte Carlo and statistical methods
PACS.64.60.Cn – Order-disorder transformations; statistical mechanics of model systems

Abstract. — The traveling salesman problem (TSP) consists of finding the length of the shortest closed tour visiting N "cities". We consider the Euclidean TSP where the cities are distributed randomly and independently in a d-dimensional unit hypercube. Working with periodic boundary conditions and inspired by a remarkable universality in the kth nearest neighbor distribution, we find for the average optimum tour length $\langle L_{\mathrm{E}} \rangle = \beta_{\mathrm{E}}(d)\, N^{1-1/d}\, [1 + O(1/N)]$ with $\beta_{\mathrm{E}}(2) = 0.7120 \pm 0.0002$ and $\beta_{\mathrm{E}}(3) = 0.6979 \pm 0.0002$. We then derive analytical predictions for these quantities using the random link approximation, where the lengths between cities are taken as independent random variables. From the "cavity" equations developed by Krauth, Mézard and Parisi, we calculate the associated random link values $\beta_{\mathrm{RL}}(d)$. For $d = 1, 2, 3$, numerical results show that the random link approximation is a good one, with a discrepancy of less than 2.1% between $\beta_{\mathrm{E}}(d)$ and $\beta_{\mathrm{RL}}(d)$. For large d, we argue that the approximation is exact up to $O(1/d^2)$ and give a conjecture for $\beta_{\mathrm{E}}(d)$, in terms of a power series in $1/d$, specifying both leading and subleading coefficients.

Résumé. — Le problème du voyageur de commerce (TSP) consiste à trouver le chemin fermé le plus court qui relie N "villes". Nous étudions le TSP euclidien où les villes sont distribuées au hasard de manière décorrélée dans l'hypercube de côté 1, en dimension d. En imposant des conditions aux bords périodiques et guidés par une universalité remarquable de la distribution des $k$ièmes voisins, nous trouvons la longueur moyenne du chemin optimal $\langle L_{\mathrm{E}} \rangle = \beta_{\mathrm{E}}(d)\, N^{1-1/d}\, [1 + O(1/N)]$, avec $\beta_{\mathrm{E}}(2) = 0{,}7120 \pm 0{,}0002$ et $\beta_{\mathrm{E}}(3) = 0{,}6979 \pm 0{,}0002$. Nous établissons ensuite des prédictions analytiques sur ces quantités à l'aide de l'approximation de liens aléatoires, où les longueurs entre les villes sont des variables aléatoires indépendantes. Grâce aux équations "cavité" développées par Krauth, Mézard et Parisi, nous obtenons dans le cas de liens aléatoires les valeurs, $\beta_{\mathrm{RL}}(d)$, analogues à $\beta_{\mathrm{E}}(d)$. Pour $d = 1, 2, 3$, les résultats numériques confirment que l'approximation de liens aléatoires est bonne, conduisant à un écart inférieur à 2,1% entre $\beta_{\mathrm{E}}(d)$ et $\beta_{\mathrm{RL}}(d)$. Pour d grand, nous donnons des arguments montrant que cette approximation est exacte jusqu'à l'ordre $1/d^2$ et nous proposons une conjecture pour $\beta_{\mathrm{E}}(d)$, exprimée en fonction d'une série en $1/d$, dont on donne les deux premiers ordres.

(*) *Present address*: W.K. Kellogg Radiation Laboratory, California Institute of Technology, Pasadena, CA 91125, USA
(**) Author for correspondence (e-mail: percus@ipno.in2p3.fr)
(***) Unité de Recherche des Universités Paris XI et Paris VI associée au CNRS

1. Introduction

Given N "cities" and the distances between them, the traveling salesman problem (TSP) consists of finding the length of the shortest closed "tour" (path) visiting every city exactly once, where the tour length is the sum of the city-to-city distances along the tour. The TSP is NP-complete, which suggests that there is no general algorithm capable of finding the optimum tour in an amount of time polynomial in N. The problem is thus simple to state, but very difficult to solve. It also happens to be the most well known combinatorial optimization problem, and has attracted interest from a wide range of fields. In operations research, mathematics and computer science, researchers have concentrated on algorithmic aspects. A particular focus has been on heuristic algorithms — algorithms which do not guarantee optimal tours — for cases where exact methods are too slow to be of use. The most effective heuristics are based on local search methods, which start with a non-optimal tour and iteratively improve the tour within a well-defined "neighborhood"; a famous example is the Lin-Kernighan heuristic [1]. More recent efforts have involved combining local search and non-deterministic methods, in order to refine heuristics to the point where they give good enough solutions for practical purposes; a powerful such technique is Chained Local Optimization [2].

Over the last fifteen years, physicists have increasingly been drawn to the TSP as well, and particularly to *stochastic* versions of the problem, where instances are randomly chosen from an ensemble. The motivation has often been to find properties applicable to a large class of disordered systems, either through good approximate methods or through exact analytical approaches. In our work, we consider two such stochastic TSPs. The first, the Euclidean TSP, is the more classic form of the problem: N cities are placed randomly and independently in a d-dimensional hypercube, and the distances between cities are defined by the Euclidean metric. The second, the random link TSP, is a related problem developed within the context of disordered systems: rather than specifying the positions of cities, we specify the lengths l_{ij} separating cities i and j, where the l_{ij} are taken to be independent, identically distributed random variables. The appeal of the random link problem is, on the one hand, that an analytical approach exists for solving it [3, 4], and on the other hand, that when certain correlations are neglected this TSP can be made to resemble the Euclidean TSP. We therefore consider the random link problem as a *random link approximation* to the (random point) Euclidean problem. Researchers outside of physics remain largely unaware of the analytical progress made on the random link TSP; one of our hopes is to demonstrate how these results are of direct interest in problems where the aim is to find the optimum Euclidean TSP tour length.

Our approach in this paper is then to examine both the Euclidean problem and the random link problem — the latter for its own theoretical interest as well as for a better understanding of the Euclidean case. We begin by considering in depth the Euclidean TSP, including a review of previous work. We find that, given periodic boundary conditions (toroidal geometry), the Euclidean optimum tour length L_{E} averaged over the ensemble of all possible instances has the finite size scaling behavior

$$\langle L_{\mathrm{E}} \rangle = \beta_{\mathrm{E}}(d)\, N^{1-1/d} \left[1 + O\left(\frac{1}{N} \right) \right] . \tag{1}$$

From simulations, we extract very precise numerical values for $\beta_{\mathrm{E}}(d)$ at $d = 2$ and $d = 3$; methodological and numerical procedures are detailed in the appendices. We also give numerical evidence that the probability distribution of L_{E} becomes Gaussian in the large N limit. In addition to these TSP results, we find a surprising universality in the scaling of the mean distance between kth nearest neighbors, for points randomly distributed in the d-dimensional hypercube. Finally, we discuss the expected behavior of $\beta_{\mathrm{E}}(d)$ in the large d limit.

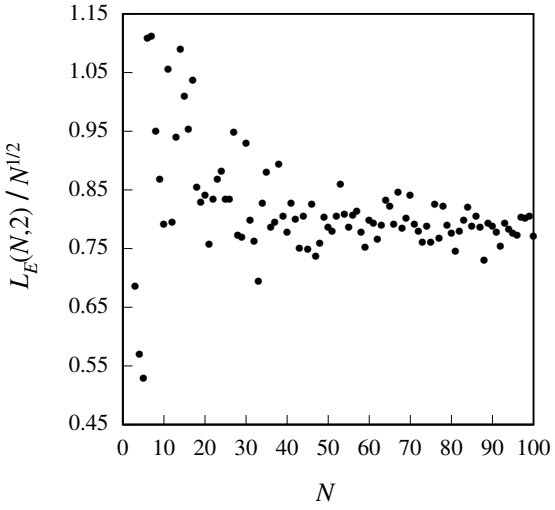

Fig. 1. — Self-averaging of 2-D Euclidean TSP optimum: convergence of $L_{\rm E}(N,2)/N^{1/2}$ on a sequence of random instances at increasing N.

In the second part of the paper we discuss the random link problem, considering it as an approximation to the Euclidean problem. Making use of the cavity method, we compare the random link $\beta_{\rm RL}(d)$ with the Euclidean $\beta_{\rm E}(d)$ values obtained from our simulations. We find that the random link approximation is correct to within 2% at $d = 2$ and 3. The rest of the section studies the large d limit of the random link model and its implications for the Euclidean TSP. We examine analytically how $\beta_{\rm RL}(d)$ scales at large d, and we relate the $1/d$ coefficient of the associated power series to an underlying d-independent "renormalized" model. Finally, we present a theoretical analysis based on the Lin-Kernighan heuristic, suggesting strongly that the relative difference between $\beta_{\rm RL}(d)$ and $\beta_{\rm E}(d)$ is positive and of $O(1/d^2)$. The random link results then lead to our large d Euclidean conjecture:

$$\beta_{\rm E}(d) = \sqrt{\frac{d}{2\pi e}}\,(\pi d)^{1/2d}\left[1 + \frac{2 - \ln 2 - 2\gamma}{d} + O\left(\frac{1}{d^2}\right)\right], \tag{2}$$

where γ is Euler's constant.

2. The Euclidean TSP

2.1. Scaling at Large N. — One of the earliest analytical results for the Euclidean TSP is due to Beardwood, Halton and Hammersley [5] (BHH). The authors considered N cities, distributed randomly and independently in a d-dimensional volume with distances between cities given by the Euclidean metric. They showed that, when the volume is the unit hypercube and the distribution of cities uniform, $L_{\rm E}/N^{1-1/d}$ is self-averaging. This means that with probability 1,

$$\lim_{N\to\infty} \frac{L_{\rm E}}{N^{1-1/d}} = \beta_{\rm E}(d), \tag{3}$$

where $\beta_{\rm E}(d)$ is independent of the randomly chosen instances. This property is illustrated in Figure 1. (In fact, the BHH result is more general than this and concerns an arbitrary volume and arbitrary form of the density of cities.) For a physics audience this large N limit

is equivalent, in appropriate units, to an infinite volume limit at constant density. $L_{\rm E}/N^{1-1/d}$ then corresponds to an energy density that is self-averaging and has a well-defined infinite volume limit. The original proof by BHH is quite complicated; simpler proofs have since been given by Karp and Steele [6, 7].

One of our goals is to determine $\beta_{\rm E}(d)$. BHH gave rigorous lower and upper bounds as a function of dimension. For any given instance, a trivial lower bound on $L_{\rm E}$ is the sum over all cities i of the distance between i and its nearest neighbor in space. In fact, since a tour at best links a city with its *two* nearest neighbors, this bound can be improved upon by summing, over all i, the mean of i's nearest and next-nearest neighbor distances. Taking the ensemble average of this quantity (that is, the average over all instances) leads to the best analytical lower bound to date. For upper bounds, BHH introduced a heuristic algorithm, now known as "strip", in order to generate near-optimal tours (discussed also in a paper by Armour and Wheeler [8]). In two dimensions the method involves dividing the square into adjacent columns or strips, and sequentially visiting the cities on a given strip according to their positions along it. The respective lower and upper bounds give $0.6250 \leq \beta_{\rm E}(2) \leq 0.9204$.

In addition to bounds, it is possible to obtain numerical *estimates* for $\beta_{\rm E}(d)$. BHH used two instances, $N = 202$ and $N = 400$, from which they estimated $\beta_{\rm E}(2) \approx 0.749$ using hand-drawn tours. Surprisingly little has been done to improve upon this value in two dimensions, and essentially nothing in higher dimensions. Stein [9] has found $\beta_{\rm E}(2) \approx 0.765$, which is frequently cited. Only recently have better values been obtained, but as they come from near-optimal tours found by heuristic algorithms, they should be considered more as upper bounds than as estimates. Using a local search heuristic known as "3-opt" [10], Ong and Huang [11] have found $\beta_{\rm E}(2) \leq 0.743$; using another heuristic, "tabu" search, Fiechter [12] has found $\beta_{\rm E}(2) \leq 0.731$; and using a variant of simulated annealing, Lee and Choi [13] have found $\beta_{\rm E}(2) \leq 0.721$. In what follows we shall show what is needed for a more precise estimate of $\beta_{\rm E}(d)$ with, furthermore, a way to quantify the associated error.

2.2. Extracting $\beta_{\rm E}(d)$. — As $N \to \infty$, $L_{\rm E}/N^{1-1/d}$ converges with probability 1 to the instance-independent $\beta_{\rm E}(d)$. Our estimate of $\beta_{\rm E}(d)$ must rest on some assumptions, though, since only finite values of N are accessible numerically. Note first that at values of N where computation times are reasonable, $L_{\rm E}$ has substantial instance-to-instance fluctuations. To reduce and at the same time quantify these fluctuations, we average over a large number of instances. We thus consider the numerical mean of $L_{\rm E}$ over the instances sampled, which itself satisfies the asymptotic relation (3) but with a smoother convergence. To extract $\beta_{\rm E}(d)$, we must understand precisely what this convergence in N is.

If cities were randomly distributed in the hypercube with open boundary conditions, the cities near the boundaries would have fewer neighbors and therefore lengthen the tour. In standard statistical mechanical systems at constant density, boundary effects lead to corrections of the form surface over volume. For the TSP at constant density, the volume grows as N and the surface as $N^{1-1/d}$. In a d-dimensional unit hypercube, then, the ensemble average of $L_{\rm E}$ would presumably have the large N behavior

$$N^{1-1/d}\,\beta_{\rm E}(d)\left(1+\frac{A}{N^{1/d}}+\frac{B}{N^{2/d}}+\cdots\right). \tag{4}$$

In order to extract $\beta_{\rm E}(d)$ numerically, it would be necessary to perform a fit which includes these corrections. A reliable numerical fit, however, must have few adjustable parameters, and the slow convergence of this series would prevent us from extracting $\beta_{\rm E}(d)$ to high accuracy. We therefore have chosen to eliminate these boundary (surface) effects by using *periodic* boundary conditions in all directions. This should not change $\beta_{\rm E}(d)$, but leaves us with fewer adjustable

parameters and a faster convergence, enabling us to work with smaller values of N where numerical simulations are not too slow.

For the hypercube with periodic boundary conditions, let us introduce the notation

$$\beta_{\mathrm{E}}(N,d) \equiv \frac{\langle L_{\mathrm{E}}(N,d)\rangle}{N^{1-1/d}}, \tag{5}$$

where $\langle L_{\mathrm{E}}\rangle$ is the average of L_{E} over the ensemble of instances. ($\beta_{\mathrm{E}}(N,d)$ is, in physical units, the zero-temperature energy density.) We then wish to understand how $\beta_{\mathrm{E}}(N,d)$ converges to its large N limit, $\beta_{\mathrm{E}}(d)$. In standard statistical mechanical systems, there is a characteristic correlation length ξ. Away from a critical point, ξ is finite, and finite size corrections decrease as $\mathrm{e}^{-W/\xi}$, where W is a measure of the system "width". At a critical point, ξ is infinite, and finite size corrections decrease as a power of $1/W$. For *disordered* statistical systems, however, this picture must be modified. Even if ξ is finite for each instance in the ensemble, the fluctuating disorder can still give rise to power-law corrections for ensemble averaged quantities. In the case of the TSP, this is particularly clear: the disorder in the positions of the cities induces large finite size effects even on simple geometric quantities.

To see how this might affect the convergence of $\beta_{\mathrm{E}}(N,d)$, consider the following. For a given configuration of N points, call $D_k(N,d)$ the distance between a point and its kth nearest neighbor, where $k = 1,\ldots,N-1$. Take the points to be distributed randomly and uniformly in the unit hypercube. Let us find $\langle D_k(N,d)\rangle$. Under periodic boundary conditions, the probability density $\rho(l)$ of finding a point at distance l from another point is simply equal (for $0 \le l \le 1/2$) to the surface area at radius l of the d-dimensional sphere:

$$\rho(l) = \frac{d\,\pi^{d/2}}{\Gamma(d/2+1)}\, l^{d-1}. \tag{6}$$

The probability of finding a point's kth nearest neighbor at distance l (see Fig. 2) is equal to the probability of finding $k-1$ (out of $N-1$) points within l, one point at l and the remaining $N-k-1$ points beyond l:

$$P[D_k(N,d)=l] = \binom{N-1}{k-1}\left[\int_0^l \rho(l')\,\mathrm{d}l'\right]^{k-1}(N-k)\;\rho(l)\left[1-\int_0^l \rho(l')\,\mathrm{d}l'\right]^{N-k-1} \tag{7}$$

$$= \binom{N-1}{k-1}(N-k)\,d\left[\frac{\pi^{d/2}}{\Gamma(d/2+1)}\right]^{k} l^{dk-1}\left[1-\frac{\pi^{d/2}}{\Gamma(d/2+1)}l^{d}\right]^{N-k-1}, \tag{8}$$

giving the ensemble average

$$\langle D_k(N,d)\rangle = \binom{N-1}{k-1}(N-k)\,d\left[\frac{\pi^{d/2}}{\Gamma(d/2+1)}\right]^{k} \times\int_0^{1/2} l^{dk}\left[1-\frac{\pi^{d/2}}{\Gamma(d/2+1)}l^{d}\right]^{N-k-1}\mathrm{d}l\;+\cdots \tag{9}$$

where the corrections are due to the $l > 1/2$ case, and are exponentially small in N.

Recognizing the integral, up to a simple change of variable, as a Beta function ($B(a,b) \equiv \int_0^1 t^{a-1}(1-t)^{b-1}\,\mathrm{d}t = \Gamma(a)\Gamma(b)/\Gamma(a+b)$) plus a further remainder term exponentially small in

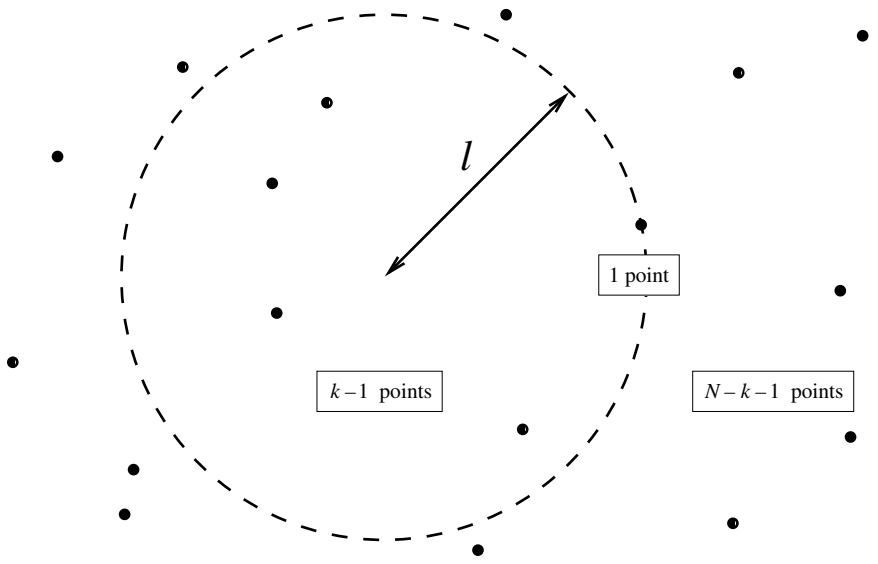

Fig. 2. — A point's $N-1$ neighbors: $k-1$ nearest neighbors are within distance l, kth nearest neighbor is at l, and remaining $N-k-1$ points are beyond l.

N, we see that

$$\langle D_k(N,d)\rangle = \frac{\Gamma(d/2+1)^{1/d}}{\sqrt{\pi}}\,\frac{\Gamma(k+1/d)}{\Gamma(k)}\,\frac{\Gamma(N)}{\Gamma(N+1/d)} + \cdots \tag{10}$$

$$= \frac{\Gamma(d/2+1)^{1/d}}{\sqrt{\pi}}\,\frac{\Gamma(k+1/d)}{\Gamma(k)}\,N^{-1/d}\left[1+\frac{1/d-1/d^2}{2N}+O\left(\frac{1}{N^2}\right)\right]. \tag{11}$$

We are confronted here with a remarkable, and hitherto unexplored, universality: the *exact* same $1/N$ series gives the N-dependence regardless of k. The same finite size scaling behavior therefore applies to all kth nearest neighbor distances.

It might be hoped then that the typical link length in optimum tours would have this N-dependence, and that $\beta_{\rm E}(N,d)$ would therefore have the same $1/N$ expansion. This is not quite the case. The link between cities i and j figures in the average $\langle D_k(N,d)\rangle$ whenever j is the kth neighbor of i; it figures in $\beta_{\rm E}(N,d)$, however, only when it belongs to the optimal tour. Two different kinds of averages are being taken, and so finite size corrections need not be identical. Nevertheless, it remains plausible that $\beta_{\rm E}(N,d)$ has a $1/N$ series expansion, albeit a different one from (11). While we cannot prove this property, it is confirmed by an analysis of our numerical data.

Our approach to finding $\beta_{\rm E}(d)$ is thus as follows: (i) we consider the ensemble average $\langle L_{\rm E}\rangle$, rather than $L_{\rm E}$ for a given instance, in order to have a quantity with a well-defined dependence on N; (ii) we use periodic boundary conditions to eliminate surface effects; (iii) we sample the ensemble using numerical simulations, and measure $\beta_{\rm E}(N,d)$ within well controlled errors; (iv) we extract $\beta_{\rm E}(d)$ by fitting these values to a $1/N$ series.

2.3. Finite Size Scaling Results. — Let us consider the $d=2$ case in detail. We found the most effective numerical optimization methods for our purposes to be the local search heuristics Lin-Kernighan (LK) [1] and Chained Local Optimization (CLO) [2] mentioned in

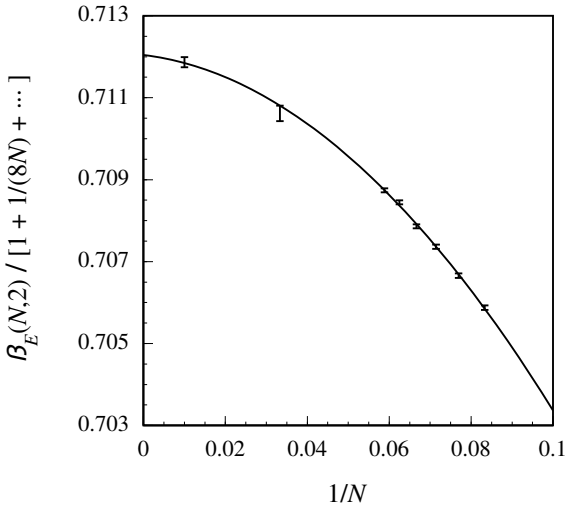

Fig. 3. — Finite size dependence of the rescaled 2-D Euclidean TSP optimum. Best fit ($\chi^2 = 5.56$) gives: $\beta_{\rm E}(N,2)/[1+1/(8N)+\cdots] = 0.7120(1-0.0171/N-1.048/N^2)$. Error bars represent statistical errors.

the introduction. Both heuristics, by definition, give tour lengths that are not always optimal. However, it is not necessary that the optimum be found 100% of the time: there is already a significant statistical error arising from instance-to-instance fluctuations, and so a further systematic error due to non-optimal tours is acceptable as long as this error is kept negligible compared to the statistical error. Our methods, along with relevant numerical details, are discussed in the appendices. For the present purposes, let us simply mention the general nature of the two heuristics used. LK works by performing a "variable-depth" local search, as discussed further in Section 3.6. CLO works by an iterative process combining LK optimizations with random perturbations to the tour, in order to explore many different local neighborhoods. We used LK for "small" N values ($N \leq 17$), averaging over 250,000 instances at each value of N, and we used CLO for "large" N values ($N = 30$ and $N = 100$), averaging over 10,000 and 6,000 instances respectively.

We fitted our resulting $\beta_{\rm E}(N,d)$ estimates to a truncated $1/N$ series: the fits are good, and are stable with respect to the use of sub-samples of the data. For a fit of the form $\beta_{\rm E}(N,d) = \beta_{\rm E}(d)(1 + A/N + B/N^2)$, we find $\beta_{\rm E}(2) = 0.7120 \pm 0.0002$, with $\chi^2 = 5.57$ for 8 data points and 3 fit parameters (5 degrees of freedom). Our error estimate for $\beta_{\rm E}(2)$ is obtained by the standard method of performing fits using a range of fixed values for this parameter: the error bar ± 0.0002 is determined by the values of $\beta_{\rm E}(2)$ which make χ^2 exceed its original result by exactly 1, *i.e.*, making $\chi^2 = 6.57$ in this case.

It is possible to extract another $\beta_{\rm E}(N,d)$ estimate by making direct use of the universality discussed previously: the universal $1/N$ series in (11) suggests that there will be a faster convergence if we use the rescaled data $\beta_{\rm E}(N,2)/[1+1/(8N)+\cdots]$. This also has the appealing property of leading to a function monotonic in N, as shown in Figure 3. We find

$$\frac{\beta_{\rm E}(N,2)}{1+1/(8N)+\cdots} \approx 0.7120\left(1-\frac{0.0171}{N}-\frac{1.048}{N^2}\right) \tag{12}$$

with the leading term having the same error bar of ± 0.0002 as before. Note that the $1/N$ term in the fit is small — 2 orders of magnitude smaller than the leading order coefficient — and so to first order the $1 + 1/8N + \cdots$ series is itself a good approximation.

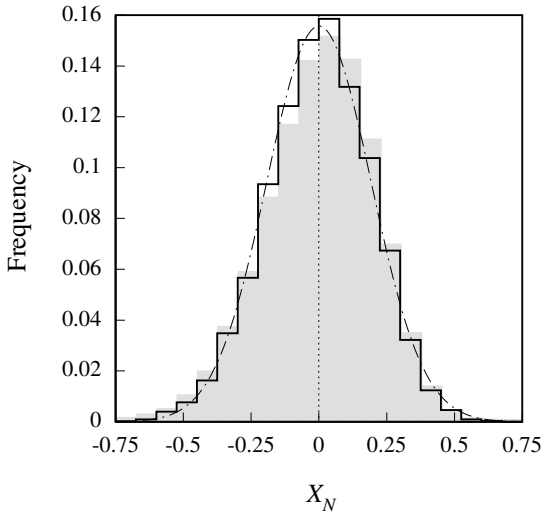

Fig. 4. — Distribution of 2-D Euclidean TSP scaling variable $X_N = (L_{\rm E} - \langle L_{\rm E}\rangle)/N^{1/2-1/d}$. Shaded region is for $N = 12$ (100,000 instances used) and solid line is for $N = 30$ (10,000 instances used). Superimposed curve shows (extrapolated) limiting Gaussian.

The same methodology was applied to the $d = 3$ case. The χ^2's again confirmed the functional form of the fit, and we find from our data $\beta_{\rm E}(3) = 0.6979 \pm 0.0002$. Also, since our initial work [14], Johnson *et al.* have performed simulations at $d = 2, 3, 4$, obtaining results [15] consistent with ours: $\beta_{\rm E}(2) \approx 0.7124$, $\beta_{\rm E}(3) \approx 0.6980$ and $\beta_{\rm E}(4) \approx 0.7234$.

2.4. Distribution of Optimum Tour Lengths. — While BHH and others [6, 7] have shown that the variance of $L_{\rm E}/N^{1-1/d}$ goes to zero as $N \to \infty$ (see also Fig. 1), they have not determined how fast this variance decreases. More generally, one might ask how the *distribution* of $L_{\rm E}/N^{1-1/d}$ behaves as $N \to \infty$. We are aware of only one result, by Rhee and Talagrand [16], showing that the probability of finding $L_{\rm E}$ with $|L_{\rm E} - \langle L_{\rm E}\rangle| > t$ is smaller than $K\exp(-t^2/K)$ for some K. Unfortunately this is not strong enough to give bounds on the variance.

Let us characterize the distribution at $d = 2$ by numerical simulation. For motivation, consider the analogy between $L_{\rm E}/N^{1-1/d}$ and E/V, the energy density in a disordered statistical system. If the system's correlation length ξ is finite (the system is not critical), E/V has a distribution which becomes Gaussian when $V \to \infty$. This is because as the subvolumes increase, the energy densities in each subvolume become uncorrelated; the central limit theorem then applies. A consequence is that σ^2, the variance of E/V, decreases as V^{-1}. If ξ is infinite (the system is critical), then in general the distribution of E/V is not Gaussian. In both cases though, the self-averaging of E/V suggests that the scaling variable $X = (E - \langle E\rangle)/\sigma V$ has a limiting distribution when $V \to \infty$.

In the case of the TSP, it can be argued using a theoretical analysis of the LK heuristic that at $d \geq 2$ the system is not critical. By analogy with E/V, if we take subvolumes to contain a fixed number of cities, the central limit theorem then suggests that $L_{\rm E}/N^{1-1/d}$ has a Gaussian distribution with σ^2 decreasing as N^{-1}. The scaling variable $X_N = (L_{\rm E} - \langle L_{\rm E}\rangle)/N^{1/2-1/d}$ should consequently have a Gaussian distribution with a finite width for $N \to \infty$ (and at $d \geq 2$). Numerical results at $d = 2$ (see Fig. 4) give good support for this.

2.5. CONJECTURES ON THE LARGE d LIMIT. — In most statistical mechanics problems, the large dimensional limit introduces simplifications because fluctuations become negligible. For the TSP, can one expect $\beta_{\rm E}(d)$ to have a simple limit as $d \to \infty$? Again, consider the property of the kth nearest neighbor distance D_k. In the large N limit, (11) gives

$$N\langle D_k(N,d)\rangle \sim N^{1-1/d}\,\frac{\Gamma(d/2+1)^{1/d}}{\sqrt{\pi}}\,\frac{\Gamma(k+1/d)}{\Gamma(k)}, \text{ or at large } d, \tag{13}$$

$$\sim N^{1-1/d}\sqrt{\frac{d}{2\pi e}}\,(\pi d)^{1/2d}\left[1+\frac{A_k}{d}+\cdots\right], \tag{14}$$

where $A_k \equiv -\gamma + \frac{1}{k-1} + \frac{1}{k-2} + \cdots$ (γ is Euler's constant). Notice that $A_k \sim \ln k$ at large k. This suggests strongly that unless the "typical" k used in the optimum tour grows exponentially in d, we may write for $d \to \infty$:

$$\beta_{\rm E}(d) = \lim_{N\to\infty} \frac{\langle L_{\rm E}(N,d)\rangle}{N^{1-1/d}} \sim \sqrt{\frac{d}{2\pi e}}\,(\pi d)^{1/2d}\left[1+O\left(\frac{1}{d}\right)\right]. \tag{15}$$

Up to $O(1/d)$, this expression is identical to the BHH lower bound on $\beta_{\rm E}(d)$ discussed in Section 2.1, given by the large N limit of $N^{1/d}\langle D_1(N,d)+D_2(N,d)\rangle/2$.

A weaker conjecture than (15) has been proposed by Bertsimas and van Ryzin [17]:

$$\beta_{\rm E}(d) \sim \sqrt{d/2\pi e} \text{ as } d \to \infty. \tag{16}$$

This limiting behavior was motivated by an analogous result for a related combinatorial optimization problem, the minimum spanning tree. Unfortunately, there is no proof of either (15) or (16); in particular, the upper bound on $\beta_{\rm E}(d)$ given by strip, discussed in Section 2.1, behaves as $\sqrt{d/6}$ at large d. Thus if the conjectures are true, the strip construction leads asymptotically to tours which are on average 1.69 times too long. Can we derive stronger upper bounds? A number of heuristic construction methods should do better than strip, but there are no reliable calculations to this effect. The only improvements over the BHH results are due to Smith [18], who generalized the strip algorithm by optimizing the shape of the strips, leading to an upper bound which is $\sqrt{2}$ times greater than the predictions of (15) and (16) at large d.

In spite of our inability to derive an upper bound which, together with the BHH lower bound, would confirm the two conjectures for $d \to \infty$, we are confident that (15) and (16) are true because of non-rigorous yet convincing arguments. One is a proof that (16) is satisfied for the TSP if it is satisfied for another related combinatorial optimization problem (see Appendix D for details). A more powerful argument, presented in Section 3.6, relies on a theoretical analysis of the LK heuristic. It suggests that up to $O(1/d^2)$, $\beta_{\rm E}(d)$ is given by a random link approximation, leading to a conjecture even stronger than (15).

3. The Random Link TSP

3.1. CORRESPONDENCE WITH THE EUCLIDEAN TSP. — Let us now consider a problem at first sight dramatically different from the Euclidean TSP. Instead of taking the positions of the N cities to be independent random variables, take the lengths $l_{ij} = l_{ji}$ between cities i and j ($1 \le i,j \le N$) to be independent random variables, identically distributed according to some $\rho(l)$. We speak of lengths rather than distances, as there is no distance metric here. This problem, introduced by physicists in the 1980s [19, 20] in search of an analytically tractable form of the traveling salesman problem, is called the *random link TSP*.

The connection between this TSP and the Euclidean TSP is not obvious, as we now have random links rather than random points. Nevertheless, one can relate the two problems. To see this, consider the probability distribution for the distance l between a fixed pair of cities (i,j) in the Euclidean TSP. This distribution, in the unit hypercube with periodic boundary conditions, is given for $0 \leq l \leq 1/2$ by the expression in (6):

$$\rho(l) = \frac{d\,\pi^{d/2}}{\Gamma(d/2+1)}\, l^{d-1}. \tag{17}$$

Of course, in the Euclidean TSP the link lengths are by no means *independent* random variables: correlations such as the triangle inequality are present. However, as noted by Mézard and Parisi [3], correlations appear exclusively when considering three or more distances, since any two Euclidean distances are necessarily independent. Let us adopt (17) as the l_{ij} distribution in the limit of small l for the random link TSP, where d in this case no longer represents physical dimension but is simply a parameter of the model. The Euclidean and random link problems then have the same small l one- and two-link distributions. In the large N limit the random link TSP may therefore be considered, rather than as a separate problem, as a *random link approximation* to the Euclidean TSP. Only joint distributions of three or more links differ between these two TSPs. If indeed the correlations involved are not too important, then the random link $\beta_{\mathrm{RL}}(d)$ can be taken as a good estimate of $\beta_{\mathrm{E}}(d)$. We shall see that this is true, particularly for large d.

3.2. Scaling at Large N. — As in the Euclidean case, we are interested in understanding the $N \to \infty$ scaling law in the random link TSP. It is relatively simple to see, following an argument similar to the one in Section 2.2, that the nearest neighbor distances D_k have a probability distribution with a scaling factor $N^{-1/d}$ at large N. Vannimenus and Mézard [20] have suggested that the random link optimum tour length with N links will then scale as $N^{1-1/d}$, and the tour will be self-averaging, *i.e.*,

$$\lim_{N\to\infty} \frac{L_{\mathrm{RL}}}{N^{1-1/d}} = \beta_{\mathrm{RL}}(d), \tag{18}$$

parallel to the BHH theorem (3) for the Euclidean case. This involves the implicit assumption that optimum tours sample a representative part of the D_k distribution, so no further N scaling effects are introduced. The assumption seems reasonable based on the analogy with the Euclidean TSP, and for our purposes we shall accept here that $\beta_{\mathrm{RL}}(d)$ exists. However, there is to our knowledge no mathematical proof of self-averaging in the random link TSP.

Following the discussion of Section 2.1, let us consider some bounds on the ensemble average $\langle L_{\mathrm{RL}} \rangle$ as derived in [20]. As before, we get a lower bound on $\beta_{\mathrm{RL}}(d)$ using nearest and next nearest neighbor distances. For an upper bound, the "strip" algorithm used in the Euclidean case (Sect. 2.1) cannot be applied to the random link case. On the other hand, Vannimenus and Mézard make use of an algorithm called "greedy" [21]: this constructs a non-optimal tour by starting at an arbitrary city, and then successively picking the link to the nearest available city until all cities are used once and a closed tour is formed. At $d > 1$, greedy gives rise to tour lengths that are self-averaging, and leads to the upper bound [20]

$$\beta_{\mathrm{RL}}(d) \leq \frac{1}{\sqrt{\pi}} \frac{\Gamma(d/2+1)^{1/d}\,\Gamma(1/d)}{d-1}. \tag{19}$$

At $d = 1$, the presumed scaling (18) suggests that $\langle L_{\mathrm{RL}} \rangle$ is independent of N, whereas greedy generates tour lengths which grow as $\ln N$. There is numerical evidence [4, 22], however, that the $d = 1$ model does indeed satisfy (18), and that $\beta_{\mathrm{RL}}(1) \approx 1.0208$.

3.3. SOLUTION *via* THE CAVITY EQUATIONS. — Since the work of Vannimenus and Mézard, several groups [23–25] have tried to "solve" the statistical mechanical problem of the random link TSP at finite temperature using the replica method, a technique developed for analyzing disordered systems such as spin glasses [26]. To date, it has only been possible to obtain part of the high temperature series of this system [23]. In view of the intractability of these replica approaches, Mézard and Parisi have derived an analytical solution using another technique from spin glass theory, the "cavity method". The details of this approach are beyond the scope of this paper, and are discussed in several technical articles [3,26,27]. For readers acquainted with the language of disordered systems, however, the broad outline is as follows: one begins with a representation of the TSP in terms of a Heisenberg (multi-dimensional spin) model in the limit where the spin dimension goes to zero. Under the assumption that this system has only one equilibrium state (no replica symmetry breaking), Mézard and Parisi have then written a recursion equation for the system when a new $(N+1)$th spin is added. The cavity method then supposes that this new spin's effect on the N other spins is negligible in the large N limit, and that its magnetization may be expressed in terms of the magnetizations of the other spins.

Using this method, Krauth and Mézard have derived a self-consistent equation for the random link TSP, at $N \to \infty$ [4]. They have determined the probability distribution of link lengths in the optimum tour in terms of $\mathcal{G}_d(x)$, where $\mathcal{G}_d(x)$ is the solution to the integral equation

$$\mathcal{G}_d(x) = \int_{-x}^{+\infty} \frac{(x+y)^{d-1}}{\Gamma(d)} \left[1 + \mathcal{G}_d(y)\right] \mathrm{e}^{-\mathcal{G}_d(y)} \, \mathrm{d}y. \tag{20}$$

Their probability distribution leads to the prediction

$$\beta_{\mathrm{RL}}(d) = \frac{d}{2\sqrt{\pi}} \left[\frac{\Gamma(d/2+1)}{\Gamma(d+1)}\right]^{1/d} \int_{-\infty}^{+\infty} \mathcal{G}_d(x) \left[1 + \mathcal{G}_d(x)\right] \mathrm{e}^{-\mathcal{G}_d(x)} \, \mathrm{d}x. \tag{21}$$

These equations can be solved numerically, as well as analytically in terms of a $1/d$ power series (see next section). At $d = 1$, Krauth and Mézard compared their prediction with the results of a direct simulation of the random link model; their numerical study [4, 22] strongly suggests that the cavity prediction is exact in this case. It has been argued, furthermore, that the cavity method is exact at $N \to \infty$ for *any* distribution of the independent random links [26]. Good numerical evidence has been found for this, notably in the case of the matching problem, a related combinatorial optimization problem [28]. The validity of the cavity assumptions therefore does not appear to be sensitive to the dimension d, and we shall assume that (21) holds for the random link TSP at all d.

Krauth and Mézard computed the $d = 1$ and $d = 2$ cases to give $\beta_{\mathrm{RL}}(1) = 1.0208$ and $\beta_{\mathrm{RL}}(2) = 0.7251$. Since $\beta_{\mathrm{RL}}(d)$ is taken to approximate $\beta_{\mathrm{E}}(d)$, let us compare these values with their Euclidean counterparts. At $d = 1$, the Euclidean TSP with periodic boundary conditions is trivial ($\beta_{\mathrm{E}}(1) = 1$); the random link TSP thus has a 2.1% relative excess. At $d = 2$, comparing with $\beta_{\mathrm{E}}(2) = 0.7120$ found in Section 2.3, the random link TSP has a 1.8% excess. In low dimensions, the random link results are then a good approximation of the Euclidean results. The approximation is better than Krauth and Mézard believed, since they made the comparison at $d = 2$ using the considerably overestimated Euclidean value of $\beta_{\mathrm{E}}(2) \approx 0.749$ from [5].

Extending the numerical solutions to higher dimensions, at $d = 3$ we find $\beta_{\mathrm{RL}}(3) = 0.7100$, which compared with $\beta_{\mathrm{E}}(3) = 0.6979$, has an excess of 1.7%. Some further random link values are $\beta_{\mathrm{RL}}(4) = 0.7322$ and $\beta_{\mathrm{RL}}(5) = 0.7639$. The value at $d = 4$ may be compared with the Euclidean estimate of Johnson *et al.* [15], $\beta_{\mathrm{E}}(4) \approx 0.7234$, giving an excess of 1.2%. The $\beta_{\mathrm{E}}(d)$ data at $d = 1, 2, 3, 4$ therefore suggest that the random link approximation improves with increasing dimension. This leads us to study the limit when d becomes large.

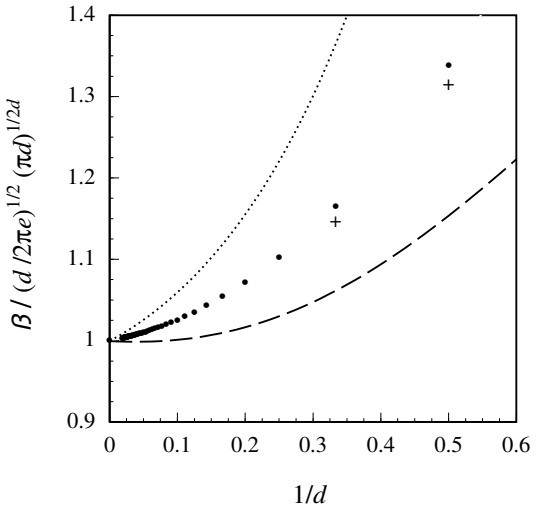

Fig. 5. — Dimensional dependence of rescaled random link TSP optimum, shown by small points, between converging "greedy" upper bound (dotted line) and nearest-neighbors lower bound (dashed line). Plus signs at $d = 2$ and $d = 3$ show Euclidean results for comparison.

3.4. Dimensional Dependence. — The large d limit was considered by Vannimenus and Mézard [20]. For $\beta_{\rm RL}(d)$, the lower bound obtained from $\langle D_1(N,d) + D_2(N,d)\rangle/2$ by way of (11) and the upper bound given in (19) differ at large d only by $O(1/d)$, giving:

$$\beta_{\rm RL}(d) = \sqrt{\frac{d}{2\pi e}}\,(\pi d)^{1/2d}\left[1 + O\left(\frac{1}{d}\right)\right]. \tag{22}$$

Note that this *exact* result is the random link analogue of the Euclidean conjecture (15).

For values of $d \leq 50$, we have calculated $\beta_{\rm RL}(d)$ numerically using the cavity equations (20,21). The results are shown in Figure 5, along with the converging upper and lower bounds, and our low d Euclidean results.

For large d, we may see whether the cavity equations are compatible with (22) by solving them analytically in terms of a $1/d$ power series. Define $\tilde{\mathcal{G}}_d(x) \equiv \mathcal{G}_d\big(\Gamma(d+1)^{1/d}\,[1/2 + x/d]\big)$. (20) may then be written:

$$\tilde{\mathcal{G}}_d(x) = \int_{-x-d}^{+\infty}\left(1 + \frac{x+y}{d}\right)^{d-1}\left[1 + \tilde{\mathcal{G}}_d(y)\right]\mathrm{e}^{-\tilde{\mathcal{G}}_d(y)}\,\mathrm{d}y \tag{23}$$

$$= \int_{-x-d}^{+\infty}\mathrm{e}^{x+y}\left[1 - \frac{1}{d}\left(x + y + \frac{(x+y)^2}{2}\right) + O\left(\frac{1}{d^2}\right)\right]\left[1 + \tilde{\mathcal{G}}_d(y)\right]\mathrm{e}^{-\tilde{\mathcal{G}}_d(y)}\,\mathrm{d}y. \tag{24}$$

Strictly speaking, the expansion of $(1 + [x+y]/d)^{d-1}$ is only valid in the interval $-x - d < y < -x + d$; however, for large y it can be shown that $\tilde{\mathcal{G}}_d(y) \sim y^d$, so the $\mathrm{e}^{-\tilde{\mathcal{G}}_d(y)}$ term in the integrand makes the $y > -x + d$ contribution exponentially small in d.

Furthermore, extending the integral's lower limit to include the region $y < -x - d$ also contributes a remainder term exponentially small in d. If we write the integral with its lower

limit at $y = -\infty$, the equation may be solved:

$$\tilde{\mathcal{G}}_d(x) = \sqrt{2}\mathrm{e}^x \left[1 - \frac{1}{d}\left(\frac{x^2}{2} + x\,\frac{3 - \ln 2 - 2\gamma}{2} - \frac{(\ln 2 + 2\gamma)^2 + 6\ln 2 + 12\gamma - 9}{8}\right) + O\left(\frac{1}{d^2}\right)\right], \tag{25}$$

where γ, we recall, represents Euler's constant. Using (21), we then find

$$\beta_{\mathrm{RL}}(d) = \sqrt{\frac{d}{2\pi e}}\,(\pi d)^{1/2d}\left[1 + \frac{2 - \ln 2 - 2\gamma}{d} + O\left(\frac{1}{d^2}\right)\right], \tag{26}$$

which is perfectly compatible with (22). This provides further evidence that the cavity method is exact for the random link TSP.

3.5. Renormalized Random Link Model at Large d. — We can motivate the large d scaling found in the previous section by examining a different sort of random link TSP. Consider a new "renormalized" model where link "lengths" x_{ij} are obtained from the original l_{ij} by the linear transformation $x_{ij} \equiv d[l_{ij} - \langle D_1(N,d)\rangle]/\langle D_1(N,d)\rangle$. Note that the x_{ij} may take on negative values, and that the nearest neighbor length in this new model has mean zero. Since the transformation is linear, there is a direct equivalence between the renormalized x_{ij} and original l_{ij} TSPs, and the two have the same optimum tours. The renormalized optimum tour length L_x may then be given in terms of the original tour length L_l by

$$L_x = d\,\frac{L_l - N\langle D_1(N,d)\rangle}{\langle D_1(N,d)\rangle}. \tag{27}$$

Now take $N \to \infty$ and $d \to \infty$. It may be seen from the l_{ij} distribution (17) and the $\langle D_1(N,d)\rangle$ expansion (14) that the random variables x_{ij} have the d-independent probability distribution $\rho(x) \sim N^{-1}\exp(x - \gamma)$. Also, in the large N limit, since L_l scales as $N^{1-1/d}$ and $\langle D_1\rangle$ scales as $N^{-1/d}$, we expect $\langle L_x\rangle \sim N\mu$ for some μ which must be, like $\rho(x)$, independent of d. Then, from (27), the TSP in the original l_{ij} variables satisfies

$$\langle L_l\rangle \sim N\langle D_1(N,d)\rangle\left[1 + \frac{\mu}{d} + O\left(\frac{1}{d^2}\right)\right], \tag{28}$$

or, using the expansion (14),

$$\beta_{\mathrm{RL}}(d) = \sqrt{\frac{d}{2\pi e}}\,(\pi d)^{1/2d}\left[1 + \frac{\mu - \gamma}{d} + O\left(\frac{1}{d^2}\right)\right]. \tag{29}$$

This result may be compared with our cavity solution of (26), where the $1/d$ coefficient is equal to $2 - \ln 2 - 2\gamma$. If the cavity method is correct at $O(1/d)$, which we strongly believe is the case, then a direct solution of the renormalized model should give $\mu = 2 - \ln 2 - \gamma$. Work is currently in progress to test this claim by numerical methods.

3.6. Large d Accuracy of the Random Link Approximation. — Since the random link model is considered to be an approximation to the Euclidean case, it is natural to ask whether the approximation becomes exact as $d \to \infty$. In this section we argue that: (i) in stochastic TSPs, good tours can be obtained using almost exclusively low order neighbors; (ii) the geometry inherent in the Euclidean TSP leads to $\beta_{\mathrm{E}}(d) \le \beta_{\mathrm{RL}}(d)$ in all dimensions d;

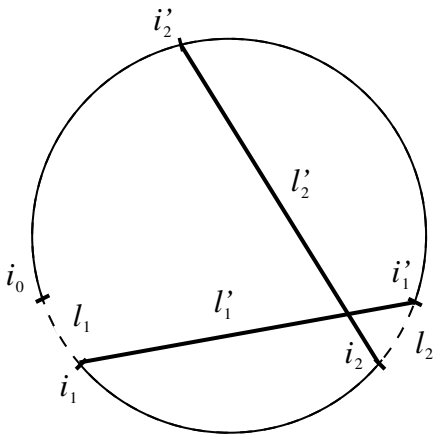

Fig. 6. — Recursive construction of removed links (dashed lines) and added links (bold lines) in an LK search.

(iii) the relative error of the random link approximation decreases as $1/d^2$ at large d. All three claims are based on a theoretical analysis of the Lin-Kernighan (LK) heuristic algorithm for constructing near-optimal tours.

The LK algorithm works as follows [1,29]. An *LK search* starts with an arbitrary tour. The principle of the search is to substitute links in the tour recursively, as illustrated schematically in Figure 6. The first step consists of choosing an arbitrary starting city i_0. Call i_1 the next city on the tour, and l_1 the link between the two. Now remove this link. Let i_1' be the nearest neighbor to i_1 that was not connected to i_1 on the original tour, and let l_1' be a new link connecting i_1 to i_1'. We now have not a tour but a "tadpole graph", containing a loop with a tail attached to it at i_1'. At this point, call i_2 one of the cities next to i_1' on the original tour, and remove the link l_2 between the two. There are two possibilities for i_2 (and thus l_2): LK chooses the one which, if we were to put in a new link between i_2 and i_0, would give a single closed tour. Now as before, let i_2' be the nearest neighbor of i_2 that was not connected to i_2 on the original tour, and let l_2' be a new link between the two. This gives a new tadpole. The process continues recursively in this manner, with the vertex hopping around while the end point stays fixed, until no new tadpoles are found. At each step, LK chooses the new i_m so as to allow the path to be closed up between i_m and i_0, forming a single tour; the result of the LK search is then the best of all such closed up tours. The *LK algorithm* consists of repeating these LK searches on different starting points i_0, each time using the current best tour as a starting tour, until no further tour improvements are possible.

Let us first sketch why the LK algorithm leads to tours which use only links between "near" neighbors, where "near" means that the neighborhood order k is small and does not grow with d. Consider any tour where a significant fraction of the links connect distant neighbors (large k). The links l_m' which the LK search substitutes for the l_m are, by definition, between very near neighbors ($k \leq 3$). As long as many long links exist, the probability at each step of substituting a near neighbor in place of a far neighbor is significant. Towards the beginning of an LK search this probability is relatively constant, so the *expected* tadpole length will decrease linearly with the number of steps. Even taking into account the fact that closing up the path between i_m and i_0 might require inserting a link with $k > 3$, there is a high probability as $N \to \infty$ that the improvement in tadpole length far outweighs this cost of closing the tour. Thus for stochastic TSPs, regardless of d, the LK algorithm can at large N replace all but a

tiny fraction of the long links with short links. It follows that in accordance with our Euclidean TSP assumption of Section 2.5, the "typical" k used in the optimum tour remains small at large d. This provides very powerful support for the $\beta_{\mathrm{E}}(d)$ conjectures (15) and (16). A consequence, making use of the exact asymptotic $\beta_{\mathrm{RL}}(d)$ result (22), is that the relative difference between $\beta_{\mathrm{E}}(d)$ and $\beta_{\mathrm{RL}}(d)$ is at most of $O(1/d)$.

Our second argument concerns why $\beta_{\mathrm{RL}}(d)$ must be greater than $\beta_{\mathrm{E}}(d)$ at all d. For the random link TSP there is no triangle inequality, which means that given two edges of a triangle, the third edge is on average longer than it would be for the Euclidean TSP. Applying this to our LK search, we can expect the link between i_m and i_0 closing up the tour to be longer in the random link case than in the Euclidean case. Thus on average, the LK algorithm will find longer random link tours than Euclidean tours. In fact, this property holds as well for any LK-like algorithm where the method of choosing the l_m and l'_m links is generalized. If the algorithm were to allow *all* possibilities for l_m and l'_m, we would be sure of obtaining the exact optimum tour, given a long enough search. In that case, the inequality on the tour lengths found by our algorithm leads directly to $\beta_{\mathrm{RL}}(d) > \beta_{\mathrm{E}}(d)$. Not surprisingly, the numerical data confirm this inequality at d up to 4 (although one should be cautious when applying the argument at $d = 1$). Note also that the inequality in itself implies conjectures (15) and (16) for the Euclidean model, since it supplies precisely the upper bound we need on $\beta_{\mathrm{E}}(d)$.

Finally let us explain why the relative difference between $\beta_{\mathrm{RL}}(d)$ and $\beta_{\mathrm{E}}(d)$ should be of $O(1/d^2)$. This involves quantifying the tour length improvement discussed above. It is clear that any non-optimal tour can be improved to the point where links are mostly between neighbors of low order. If LK, or a generalized LK-like algorithm, is able to improve the tour further, the relative difference in length will be of $O(1/d)$; we see this from (14), noting that the neighborhood order k is small both before and after the LK search. Now we need to quantify the *probability* that LK indeed succeeds in improving the tour. We may consider the vertex of the LK tadpole graph as executing a random walk, in which case the probability of closing up a tour by a sufficiently short link is equivalent to the probability of the random walk's end-to-end distance being sufficiently small. In that case it may be shown that, over the course of an LK search, the probability of successfully closing a random link tour minus the probability of successfully closing a Euclidean tour scales at large d as $2/(d-2)$. From this, we conclude that improvements in the Euclidean model are $O(1/d)$ more probable than in the random link model. Now, the relative tour length improvement for the Euclidean TSP compared to the random link TSP is simply the relative tour length improvement *when a better tour is found*, times the probability of finding a better tour — hence $O(1/d^2)$. If we consider a generalized LK search as described in the previous paragraph, where the algorithm necessarily finds the true optimum, then this result applies to the exact β's: the relative difference between $\beta_{\mathrm{RL}}(d)$ and $\beta_{\mathrm{E}}(d)$ will scale at large d as $1/d^2$.

Three comments are in order concerning this surprisingly good accuracy of the random link approximation. First, the factor $2/(d-2)$ is only appropriate for large d. It is not small even for $d = 4$. (Its divergence at $d = 2$ is associated with the fact that a two-dimensional random walk returns to its origin with probability 1.) We therefore expect the $1/d^2$ scaling to become apparent only for $d \geq 5$, beyond the range of our numerical data. Second, we have seen that the coefficient of the $1/d$ term in $\beta_{\mathrm{RL}}(d)$ may be obtained by the cavity method. Assuming that this method is correct and that $\beta_{\mathrm{RL}}(d)$ and $\beta_{\mathrm{E}}(d)$ do indeed converge as $1/d^2$, this leads to a particularly strong conjecture for the Euclidean TSP:

$$\beta_{\mathrm{E}}(d) = \sqrt{\frac{d}{2\pi e}}\,(\pi d)^{1/2d}\left[1 + \frac{2 - \ln 2 - 2\gamma}{d} + O\left(\frac{1}{d^2}\right)\right]. \tag{30}$$

Third, this type of LK analysis can in fact be extended to many other combinatorial optimization problems, such as the assignment, matching and bipartite matching problems. In these cases, we expect the random link approximation to give rise to a $O(1/d^2)$ relative error just as in the TSP.

4. Summary and Conclusions

The first goal in our work has been to investigate the finite size scaling of $L_{\rm E}$, the optimum Euclidean traveling salesman tour length, and to obtain precise estimates for its large N behavior. Motivated by a remarkable universality in the kth nearest neighbor distribution, we have found that under periodic boundary conditions, the convergence of $\langle L_{\rm E}\rangle/N^{1-1/d}$ to its limit $\beta_{\rm E}(d)$ is described by a series in $1/N$. This has enabled us to extract $\beta_{\rm E}(2)$ and $\beta_{\rm E}(3)$ using numerical simulations at small values of N, where errors are easy to control. Furthermore, thanks to a bias-free variance reduction method (see Appendix B), these estimates are extremely precise.

Our second goal has been to examine the random link TSP, where there are no correlations between link lengths. We have considered it as an approximation to the Euclidean TSP, in order to understand better the dimensional scaling of $\beta_{\rm E}(d)$. For small d, we have used the cavity method to obtain numerical values of the random link $\beta_{\rm RL}(d)$. Comparing these with our numerical values for $\beta_{\rm E}(d)$ shows that the random link approximation is remarkably good, accurate to within 2% at low dimension. For large d, we have solved the cavity equations analytically to give $\beta_{\rm RL}(d)$ in terms of a $1/d$ series. We have then argued, using a theoretical analysis of iterative tour improvement algorithms, that the relative difference between $\beta_{\rm RL}(d)$ and $\beta_{\rm E}(d)$ decreases as $1/d^2$. This leads to our conjecture (30) on the large d behavior of $\beta_{\rm E}(d)$, specifying both its asymptotic form and its leading order correction.

Let us conclude with some remaining open questions. First of all, while the cavity method most likely gives the exact result for the random link TSP, we would be interested in seeing this argued on a more fundamental physical level. Readers with a background in disordered systems will recognize that the underlying assumption of a unique equilibrium state is false in many NP-complete problems, and in particular in the spin-glass problem that has inspired the cavity method. What makes the TSP different? Second of all, our renormalized random link model provides an alternate approach to finding the $1/d$ coefficient of the power series in $\beta_{\rm RL}(d)$, and could prove a useful test of the cavity method's validity. A solution to the renormalized model using heuristic methods appears within reach. Third of all, the $O(1/d^2)$ convergence of the random link approximation merits further study, from both numerical and analytical perspectives. Numerically, Euclidean simulations at $d \geq 5$ could provide powerful support for the form of the convergence, and thus for our conjecture (30). Analytically, the qualitative arguments presented in Section 3.6, based on the LK algorithm, could perhaps be refined by a more quantitative approach. Lastly, it is worth noting that the $O(1/d^2)$ convergence should apply equally well to the *distribution* of link lengths in the optimum tour. The random link prediction for this distribution can be obtained from the cavity method [4]; an interesting test would then be to compare it with simulation results for the true Euclidean distribution.

Acknowledgments

We thank E. Bogomolny, D.S. Johnson, M. Mézard, S.W. Otto and N. Sourlas for fruitful discussions. NJC acknowledges a grant from the Human Capital and Mobility Program of the Commission of the European Communities; JBdM acknowledges financial support from the French Ministry of Higher Education and Research; OCM acknowledges support from NATO travel grant CRG 920831 and from the Institut Universitaire de France.

Appendix A

Overview of the Numerical Methodology

In the following, we discuss the procedures used to obtain the raw data from which $\beta_{\rm E}(d)$ and the finite size scaling coefficients are extracted. Two major problems must be solved in order to get good estimates of $\beta_{\rm E}(N,d)$. First, $\beta_{\rm E}(N,d)$ is defined as an ensemble average $\langle L_{\rm E}(N,d)\rangle/N^{1-1/d}$, but is measured by a numerical average over a finite sample of instances. The instance-to-instance fluctuations in $L_{\rm E}$ give rise to a statistical error, which decreases only as the inverse square root of the sample size. Keeping the statistical error down to acceptable levels could require inordinate amounts of computing time. We therefore find it useful to introduce a variance reduction trick: instead of measuring $L_{\rm E}$, we measure $L_{\rm E}-\lambda L^*$, where λ is a free parameter and L^* can be any quantity which is strongly correlated with $L_{\rm E}$. Details are given in Appendix B.

A second and more basic problem is that it is computationally costly to determine the optimal tour lengths for a large number of instances, precisely because the TSP is an NP-complete problem. The most sophisticated "branch and cut" algorithms can take minutes on a workstation to solve a single instance of size $N \leq 100$ to optimality. However, we do not need to guarantee optimality: the statistical error in $\beta_{\rm E}(N,d)$ already limits the quality of our estimate, and so an additional (systematic) error in $L_{\rm E}$ is admissible as long as it is negligible compared to the statistical error. We may thus use fast heuristics to measure $L_{\rm E}$, rather than exact but slower algorithms. This is discussed further in Appendix C.

Appendix B

Statistical Errors and a Variance Reduction Trick

Consider estimating $\langle L_{\rm E}(N,d)\rangle$ at a given N by sampling over many instances. If we have M independent instances, the simplest estimator for $\langle L_{\rm E}(N,d)\rangle$ is $\overline{L_{\rm E}(N,d)}$, the numerical average over the M instances of the minimum tour lengths. This estimator has an expected statistical error $\sigma(M)=\sigma_{L_{\rm E}}/\sqrt{M}$, where $\sigma_{L_{\rm E}}$ is the instance-to-instance standard deviation of $L_{\rm E}$.

Now let us define L_k to be the sum, over all cities, of kth nearest neighbor distances. $\langle L_k\rangle$ is its ensemble average; in terms of the notation used earlier in the text, $\langle L_k\rangle = N\langle D_k\rangle$. It has been noted by Sourlas [30] that $L_{\rm E}$ is strongly correlated with L_1, L_2 and L_3. He therefore suggested reducing the statistical error in $\langle L_{\rm E}\rangle$ using the estimator

$$E_{\rm S} = \langle L_{123}\rangle\,\overline{L_{\rm E}/L_{123}}, \tag{B.1}$$

where L_{123} is the arithmetic mean of L_1, L_2 and L_3. The ensemble average $\langle L_{123}\rangle$ can be calculated analytically from (11), and so the variance of $E_{\rm S}$ comes from fluctuations in the ratio $L_{\rm E}/L_{123}$. If $L_{\rm E}$ were a constant factor times L_{123}, this estimator would of course be perfect, *i.e.*, it would have zero variance. This is not the case, however, and furthermore the use of a ratio biases the Sourlas estimator: its true mathematical expectation value differs from $\langle L_{\rm E}(N,d)\rangle$ by $O(1/N)$. To improve upon this, we have introduced our own bias-free estimator [31]:

$$E_{\text{M-P}} = \lambda\langle L_{12}\rangle + \overline{L_{\rm E}-\lambda L_{12}}, \tag{B.2}$$

where L_{12} is the arithmetic mean of L_1 and L_2, and λ is a free parameter. Our estimator has a reduced variance because $L_{\rm E}$ and L_{12} are correlated. It is easy to show that the variance

of $E_{\text{M-P}}$ is minimized at a unique value of λ, $\lambda^* = C(L_{\text{E}}, L_{12})\,\sigma_{L_{\text{E}}}/\sigma_{L_{12}}$, where $C(A,B) \equiv \big\langle (A - \langle A\rangle)\,(B - \langle B\rangle)\big\rangle/\sigma_A\sigma_B$ is the correlation coefficient of A and B. The variance then becomes $\sigma^2_{L_{\text{M-P}}} = \sigma^2_{L_{\text{E}}}[1 - C^2(L_{\text{E}}, L_{12})]/M$. Empirically, we have found this variance reduction procedure to be quite effective, since $\sqrt{1-C^2} \approx 0.38$ at $d = 2$ and $\sqrt{1-C^2} \approx 0.31$ at $d = 3$. The statistical error is thus reduced by about a factor of 3; this means that for a given error, computing time is reduced by about a factor of 10.

Appendix C

Control of Systematic Errors

Our procedure for estimating L_{E} at a given instance involves running a good heuristic m times from random starts on that instance, and taking the best tour length found in those m trials. The expected systematic error can be found from the frequencies with which each local optimum appears in a large number of test trials. (This large number must be much greater than m, the actual number of trials used in production runs.) The measurement is performed on a sufficiently large sample of instances, from which we extract the *average* size of the systematic error in $\langle L_{\text{E}}(N,d)\rangle$ as a function of m. We have found that in practice, this error is dominated by those infrequent instances where a sub-optimal tour is obtained with the highest frequency.

As N increases, the probability of not finding the true optimum increases rather fast; for a given heuristic, it is thus necessary to increase m with N in such a way that the systematic error remains much smaller than the statistical error. If the heuristic is not powerful enough, m will be too large for the computational resources. For our purposes, we have found that the Lin-Kernighan heuristic [1] is powerful enough for the smaller values of N ($N \leq 17$). For $20 \leq N \leq 100$, it was more efficient to switch to Chained Local Optimization (CLO) [2, 32], a more powerful heuristic which can be thought of as a generalization of simulated annealing. (When the temperature parameter is set to zero so that no up-hill moves are accepted, as was the case for our runs, CLO with embedded Lin-Kernighan is called "Iterated Lin-Kernighan" [33, 34].) With these choices, using in two dimensions $m = 10$ for $N \leq 17$ (LK), $m = 5$ for $N = 30$ and $m = 20$ for $N = 100$ (CLO), we have kept systematic errors to under 10% of the statistical errors.

Appendix D

Bounding $\beta_{\text{E}}(d)$ using the Bipartite Matching Problem

Given two sets of N points $\text{P}_1, \ldots, \text{P}_N$ and $\text{Q}_1, \ldots, \text{Q}_N$ in d-dimensional Euclidean space, the bipartite matching (BM) problem asks for the minimum matching cost L_{BM} between the P_i's and the Q_i's, with the constraint that only links of the form $\text{P} - \text{Q}$ are allowed. The cost of a matching is equal to the sum of the distances between matched pairs of points. When points P_i and Q_i are chosen at random in a d-dimensional unit hypercube, it is natural to expect $L_{\text{BM}}/N^{1-1/d}$ to be self-averaging as $N \to \infty$. To date, a proof of this property has not been given, even though the self-averaging of the analogous quantity in the more general matching problem (where links $\text{P} - \text{P}$ and $\text{Q} - \text{Q}$ are allowed as well) can be shown at all d in essentially the same way as for the TSP, following arguments developed by Steele [7]. For $d = 1$, it is in fact known that self-averaging *fails* in the BM. For large d, however, let us assume that $L_{\text{BM}}/N^{1-1/d}$ does converge to some $\beta_{\text{BM}}(d)$ in the large N limit.

We shall now derive a bound for the Euclidean TSP constant $\beta_{\rm E}(d)$ in terms of $\beta_{\rm BM}(d)$. Consider K disjoint sets $\mathrm{S}_1, \ldots, \mathrm{S}_K$, together forming a large set $\mathrm{S} \equiv \mathrm{S}_1 \cup \cdots \cup \mathrm{S}_K$, and let each S_i contain N random points in the d-dimensional unit hypercube. Construct the K minimum matchings $\mathrm{S}_1 - \mathrm{S}_2, \mathrm{S}_2 - \mathrm{S}_3, \ldots, \mathrm{S}_{K-1} - \mathrm{S}_K$ and $\mathrm{S}_K - \mathrm{S}_1$. Starting at any point in S_1, generate a loop (a closed path) in S by following the matchings $\mathrm{S}_1 - \mathrm{S}_2$, $\mathrm{S}_2 - \mathrm{S}_3$, ... until the path returns to its starting point. The set of all such distinct loops $\Omega_1, \ldots, \Omega_M$ $(M \leq N)$ is then equivalent to the set S, and furthermore the sum of the loop lengths is equal to the sum of all minimum matchings costs $(L_{\rm BM})_{\mathrm{S}_i - \mathrm{S}_{i+1}}$. (Note that $(L_{\rm BM})_{\mathrm{S}_K - \mathrm{S}_{K+1}}$ is defined as $(L_{\rm BM})_{\mathrm{S}_K - \mathrm{S}_1}$.)

Now, consider the optimum TSP tour through all the points of S_1. Construct a giant closed path visiting every point in S at least once, by substituting into this TSP tour the loops $\Omega_1, \ldots, \Omega_M$ in place of their starting points in S_1. Using standard techniques [6], we can construct from this path of length $(K+1)N$ a shorter closed path of length KN which visits every point in S exactly once. For the Euclidean TSP tour length $L_{\rm E}$, we then obtain the inequality

$$(L_{\rm E})_{\rm S} \leq (L_{\rm E})_{\mathrm{S}_1} + \sum_{i=1}^{K} (L_{\rm BM})_{\mathrm{S}_i - \mathrm{S}_{i+1}}. \tag{D.1}$$

If S consists of random points chosen independently and uniformly in the unit hypercube, then averaging over all configurations, dividing by $N^{1-1/d}$ and taking the limit $N \to \infty$, we find

$$K^{1-1/d} \beta_{\rm E}(d) \leq \beta_{\rm E}(d) + K \beta_{\rm BM}(d). \tag{D.2}$$

Letting $K = d$, this gives in the large d limit $\beta_{\rm E}(d) \leq \beta_{\rm BM}(d)$. Based on analogies with other combinatorial optimization problems [17], $\beta_{\rm BM}(d)$ is expected to scale as $\sqrt{d/2\pi e}$ when $d \to \infty$. In that case, $\beta_{\rm E}(d)$ too must satisfy the Bertsimas-van Ryzin conjecture (16).

References

[1] Lin S. and Kernighan B., *Operations Res.* **21** (1973) 498–516.
[2] Martin O.C. and Otto S.W., *Ann. Operations Res.* **63** (1996) 57–75.
[3] Mézard M. and Parisi G., *Europhys. Lett.* **2** (1986) 913–918.
[4] Krauth W. and Mézard M., *Europhys. Lett.* **8** (1989) 213–218.
[5] Beardwood J., Halton J.H. and Hammersley J.M., *Proc. Cambridge Philos. Soc.* **55** (1959) 299–327.
[6] Karp R.M. and Steele M., in The Traveling Salesman Problem, E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan and D.B. Shmoys, Eds. (John Wiley and Sons, New York, 1985).
[7] Steele M., *Ann. Probability* **9** (1981) 365–376.
[8] Armour R.S. and Wheeler J.A., *Am. J. Phys.* **51** (1983) 405–406.
[9] Stein D., Ph.D. thesis, Harvard University (1977).
[10] Lin S., *Bell Syst. Tech. J.* **44** (1965) 2245–2269.
[11] Ong H.L. and Huang H.C., *Eur. J. Operational Res.* **43** (1989) 231–238.
[12] Fiechter C.N., *Discrete Appl. Math.* **1** (1994) 243–267.
[13] Lee J. and Choi M.Y., *Phys. Rev. E* **50** (1994) R651–R654.
[14] Percus A.G. and Martin O.C., *Phys. Rev. Lett.* **76** (1996) 1188–1191.

[15] Johnson D.S., McGeoch L.A. and Rothberg E.E., "Asymptotic Experimental Analysis for the Held-Karp Traveling Salesman Bound", 7th Annual ACM-SIAM Symposium on Discrete Algorithms (Atlanta, GA, 1996) pp. 341-350.
[16] Rhee W.T. and Talagrand M., *Ann. Probability* **17** (1989) 1–8.
[17] Bertsimas D.J. and van Ryzin G., *Operations Res. Lett.* **9** (1990) 223–231.
[18] Smith W.D., Ph.D. thesis, Princeton University (1989).
[19] Kirkpatrick S. and Toulouse G., *J. Phys. France* **46** (1985) 1277–1292.
[20] Vannimenus J. and Mézard M., *J. Phys. Lett. France* **45** (1984) L1145–L1153.
[21] Papadimitriou C.H. and Steiglitz K., Combinatorial Optimization: Algorithms and Complexity (Prentice Hall, Englewood Cliffs, NJ, 1982).
[22] Krauth W., Ph.D. thesis, Université Paris-Sud (1989).
[23] Mézard M. and Parisi G., *J. Phys. Lett. France* **46** (1985) L771–L778.
[24] Orland H., *J. Phys. Lett. France* **46** (1985) L763–L770.
[25] Baskaran G., Fu Y. and Anderson P.W., *J. Stat. Phys.* **45** (1986) 1–25.
[26] Mézard M., Parisi G. and Virasoro M.A., Eds., Spin Glass Theory and Beyond (World Scientific, Singapore, 1987).
[27] Mézard M., Parisi G. and Virasoro M.A., *Europhys. Lett.* **1** (1986) 77–82.
[28] Brunetti R., Krauth W., Mézard M. and Parisi G., *Europhys. Lett.* **14** (1991) 295–301.
[29] Johnson D.S. and McGeoch L.A., in Local Search in Combinatorial Optimization, E.H.L. Aarts and J.K. Lenstra, Eds. (John Wiley and Sons, New York, 1997) to appear.
[30] Sourlas N., *Europhys. Lett.* **2** (1986) 919–923.
[31] Martin O.C. and Percus A.G., *Eur. J. Operational Res.* (1996), submitted.
[32] Martin O., Otto S.W. and Felten E.W., *Complex Systems* **5** (1991) 299–326.
[33] Johnson D.S., in Proceedings of the 17th Colloquium on Automata, Languages, and Programming (Springer-Verlag, Berlin, 1990) pp. 446–461.
[34] Martin O., Otto S.W. and Felten E.W., *Operations Res. Lett.* **11** (1992) 219–224.

Chapter III

The random link TSP and its analytical solution

Contents

Chapter overview 49
Abstract 53
1. Introduction 55
2. Background and the cavity method 56
3. Numerical analysis: $d = 2$ case 60
4. Numerical analysis: renormalized model 64
5. Conclusion 67
Bibliography 68

– *Chapter overview* –

Thus far, we have used the Euclidean stochastic TSP as the basis for discussion. We have studied the behavior of the optimal tour length $L_E(N, d)$ at large N using numerical techniques. We have seen how, when $N \to \infty$, $L_E(N, d)/N^{1-1/d}$ converges to its instance-independent limit $\beta_E(d)$. We have then taken advantage of the random link approximation to obtain $\beta_{RL}(d)$, a theoretical prediction for $\beta_E(d)$. This approximation involves substituting a different model for the Euclidean model: the *random link* TSP, in which the independent random variables are not the positions of cities but rather the lengths separating pairs of cities.

The random link TSP was developed by theoreticians in search of an analytically tractable version of the problem. A major breakthrough occurred with the use of the *cavity method*, by Mézard & Parisi (1986*b*) and later Krauth & Mézard (1989), to provide an analytical prediction for the (random link) quantity $\beta_{RL}(d)$. In this chapter, we discuss the analytical approach afforded by the cavity method. The approach relies on certain assumptions which, while plausible, have thus far been tested numerically only for a uniform distribution of link lengths — corresponding to the $d = 1$ case. We therefore perform numerical simulations to test the cavity predictions more broadly, and find convincing evidence that these predictions are correct for all d.

The strategy for attacking the random link TSP analytically (Orland, 1985) has been to map the problem onto a model with m-component spins; taking the limit $m \to 0$ leads to

a representation of a self-avoiding walk. This method, originally developed in the context of polymer theory (De Gennes, 1972), then allows us to recover the partition function for the TSP in terms of a model that is more easily solvable. The first success in solving it was due to Mézard & Parisi (1986*a*), by means of the "replica" method developed in spin glasses (see Mézard, Parisi & Virasoro, 1987). The idea behind the replica method is a simple one. In order to compute the free energy $F = -T \ln Z$ averaged over the ensemble, we imagine n independent copies (replicas) of the system, and then write

$$\langle \ln Z \rangle = \lim_{n \to 0} \frac{\langle Z^n \rangle - 1}{n}. \tag{III.1}$$

Making the assumption that all n replicas are interchangeable, *i.e.*, they all reach the same equilibrium state (this ergodicity hypothesis is known as *replica symmetry*), Mézard and Parisi obtained a high-temperature expansion for $\beta_{RL}(d = 1)$.

Unfortunately, the replica method allowed estimating the solution at $T = 0$ (the global optimum) only by a numerical extrapolation from higher temperatures. Mézard and Parisi therefore instead tried using the cavity approach, a mean-field method (also assuming replica symmetry) whereby one adds an additional spin to the system and calculates the magnetization at this site in the absence of correlations between the other spins. This mean-field assumption is generally held to be correct in the limit of a large number of spins. Krauth & Mézard (1989) then completed the resulting calculation, finding an integral equation for $\beta_{RL}(d)$. Cerf, Boutet de Monvel, Bohigas, Martin & Percus (1997) solved the integral equation numerically for small values of d, as well as in a large d expansion:

$$\beta_{RL}(d) = \sqrt{\frac{d}{2\pi e}}\,(\pi d)^{1/2d} \left[1 + \frac{2 - \ln 2 - 2\gamma}{d} + O\left(\frac{1}{d^2}\right)\right], \tag{III.2}$$

where γ is Euler's constant.

The problematic assumption used here, however, is that of replica symmetry. In the spin glass model of Sherrington & Kirkpatrick (1975), where similar methods were first tried, it was found that the replica symmetry assumption led to a ground-state energy that was off by about 5% (Mézard, Parisi & Virasoro, 1987, pp. 13–14). It was therefore necessary to resort to a *replica symmetry breaking* (RSB) scheme in order to obtain more realistic results. This involved defining an "overlap" q_{ab} between two states a and b, so as to provide a measure of "distance" between them in state space. For the TSP, for example, Kirkpatrick & Toulouse (1985) have defined q_{ab} as the proportion of links in common between two tours a and b. RSB implies, among other things, that q is not a self-averaging quantity, *i.e.*, in the large N limit the distribution $P(q)$ does not simply approach a delta function.

In the case of the random link TSP, Krauth & Mézard (1989) performed a number of $d = 1$ numerical checks indicating that, unlike for the spin glass, the replica symmetric solution given by the cavity method is correct. They solved the cavity equation numerically to obtain $\beta_{RL}(1) = 1.0208$, finding it in close agreement with the earlier direct simulations of Kirkpatrick and Toulouse — whose results suggested $\beta_{RL}(1) \approx 1.045$ — as well as with Mézard and Parisi's extrapolated value of $\beta_{RL}(1) = 1.04 \pm 0.015$ from the replica method. Recent simulations by

Johnson, McGeoch & Rothberg (1996) give $\beta_{RL}(1) = 1.0209 \pm 0.0002$ (see Appendix E), providing excellent confirmation of the theoretical predictions. Krauth and Mézard also performed simulations of their own, and compared the cavity and simulated values for the probability distribution $\mathcal{P}_d(l)$ of link lengths l in the optimal tour, finding close agreement.

More direct evidence of the lack of RSB comes from the analysis by Sourlas (1986), who considered the low-temperature statistical mechanics of the $d = 1$ random link TSP. Performing numerical simulations at temperatures as low as $T = 0.85$, he found that the distribution for the overlap function, as defined above, does indeed approach a delta function at large N. This sort of behavior is inconsistent with RSB, and suggests that the replica symmetric solution used in the cavity method is indeed exact. As Sourlas noted, however, obtaining good enough statistics for simulations is extremely difficult for lower T, where fluctuations in q become much larger.

We wish to extend these numerical confirmations of the cavity method to $d > 1$, in order to determine its validity for choices of d relevant to the Euclidean TSP. Since we are interested in $T = 0$ properties, direct analysis of the overlap function does not appear within reach. However, the sort of tests proposed by Krauth and Mézard is easily applicable to cases of interest such as $d = 2$, and another method can be used to test the $O(1/d)$ coefficient in the large d expansion (III.2) for $\beta_{RL}(d)$. (The leading order term has already been shown exact by Vannimenus & Mézard (1984).) This latter method involves defining a "renormalized" random link TSP without the parameter d, having tour lengths whose leading order behavior at large N goes as $L(N) \sim \mu N$. The quantity μ may be expressed in terms of the $O(1/d)$ coefficient in (III.2); by performing numerical simulations on this renormalized model, we manage to verify the subleading behavior of the cavity method's $\beta_{RL}(d)$ prediction.

Our simulation methods are summarized in Appendix D and follow those of Chapter II: we use heuristic algorithms at values of N up to $N = 100$, where both statistical and systematic errors can be well controlled. We extrapolate to the large N limit by fitting to the expected finite size scaling law. The resulting fits are excellent, and confirm the cavity predictions both of $\beta_{RL}(2)$ and of the subleading term in the large d $\beta_{RL}(d)$ expansion to within well under 1%. (This may be compared with the 5% error of the replica symmetric solution in the Sherrington-Kirkpatrick spin glass model.) Taken together, then, this provides good arguments for the claim that the cavity method is exact for all d, and further reason for conjecturing that there is no RSB down to $T = 0$.

The stochastic traveling salesman problem: Finite size scaling and the cavity prediction

Allon G. Percus and Olivier C. Martin
Division de Physique Théorique*
Institut de Physique Nucléaire
Université Paris-Sud
F-91406 Orsay Cedex, France.

percus@ipno.in2p3.fr
martino@ipno.in2p3.fr

Abstract

We study the random link TSP, where lengths l_{ij} between city i and city j $(i < j)$ are taken to be independent, identically distributed random variables. We discuss a theoretical approach, the cavity method, that has been proposed for finding the optimal tour length over this random ensemble, given the assumption of replica symmetry. Using finite size scaling and a renormalized model, we test the cavity predictions against the results of simulations, and find excellent agreement over a range of distributions. In doing so, we provide numerical evidence that the replica symmetric solution to this problem is the correct one.

* Unité de recherche des Universités de Paris XI et Paris VI associée au CNRS

— 1 — Introduction

Over the past 15 years, the study of the traveling salesman problem (TSP) from the point of view of statistical physics has been gaining added currency, as theoreticians have improved their understanding of the links between combinatorial optimization and disordered systems. One particular breakthrough occurred with the idea, first formulated by MÉZARD & PARISI (1986*b*) and later developed by KRAUTH & MÉZARD (1989), that a version of the stochastic TSP could be solved using an analytical method inspired from spin glasses. This method, known as the *cavity method*, is based on certain assumptions pertaining to the physical properties of the system, notably that of replica symmetry. Although in the case of the spin glass replica symmetry is known to be broken (MÉZARD, PARISI & VIRASORO, 1987), for the TSP there are various grounds for suspecting that the assumption is valid (SOURLAS, 1986; KRAUTH & MÉZARD, 1989).

The traveling salesman problem is as follows: given N sites (or "cities"), find the length of the shortest closed path ("tour") passing through all cities exactly once. In the stochastic TSP, the matrix of distances separating pairs of cities is drawn randomly from an ensemble. The ensemble that has received the most attention among theoreticians is the *random link* case, where the lengths l_{ij} between city i and city j ($i < j$) are taken to be independent random variables, all identically distributed according to some $\rho(l)$. The idea of looking at this random link ensemble, rather than the more traditional "random point" ensemble with cities distributed uniformly in Euclidean space, originated with an attempt by Kirkpatrick to find some sort of TSP equivalent to the earlier SHERRINGTON & KIRKPATRICK (1975) model for spin glasses.

The great advantage of the random link TSP over the (random point) Euclidean TSP is that with the former it is actually possible to make analytical progress on the problem. The cavity solution of MÉZARD & PARISI (1986*b*) and KRAUTH & MÉZARD (1989) leads to a system of integral equations, which can be solved — numerically at least — to give the optimal tour length in the limit $N \to \infty$.

In a previous paper (CERF, BOUTET DE MONVEL, BOHIGAS, MARTIN & PERCUS, 1997), we chose the random link distribution $\rho(l)$ to match that of the distribution of individual city-to-city distances in the Euclidean case, and used the random link TSP as a *random link approximation* to the Euclidean TSP. The approximation might seem somewhat crude, since it neglects all correlations between Euclidean distances, such as the triangle inequality. Nevertheless, it gives remarkably good results. For the d-dimensional Euclidean case in $d = 2$ and $d = 3$, a numerical solution of the (random link) cavity equations predicts asymptotic $N \to \infty$ optimal lengths within 2% of the values obtained from direct (Euclidean) simulation. In the limit $d \to \infty$, this gap shows all signs of disappearing. The random link problem is thus more closely related to the Euclidean problem than most researchers have suspected.

The random link TSP is also interesting in itself, however, particularly because little numerical work has accompanied the analytical progress made. This is all the more important given the questions surrounding the hypotheses adopted in the cavity method. In this paper we attempt to redress this imbalance, providing first of all a numerical study of the finite size scaling of the random link optimal tour length, and second of all, empirical arguments suggesting that the

cavity solution is in fact the correct one.

– 2 – Background and the cavity method

In an attempt to apply tools from statistical mechanics to optimization problems, KIRKPATRICK & TOULOUSE (1985) introduced a particularly simple case of the random link TSP. The distribution of link lengths l_{ij} was taken to be uniform, so that $\rho(l)$ is constant over a fixed interval. In light of the random link approximation, one may think of this as corresponding, at large N, to the 1-D Euclidean case. (When cities are randomly and uniformly distributed on a line segment, the distribution of lengths between pairs is uniform.) Although the 1-D Euclidean case is trivial — particularly if we adopt periodic boundary conditions, in which case the optimal tour length is simply the length of the line segment — the corresponding random link problem is far from trivial.

The simulations performed by Kirkpatrick and Toulouse suggested a random link optimal tour length value of $L_{RL} \approx 1.045$ in the $N \to \infty$ limit.[1] MÉZARD & PARISI (1986*a*) attempted to improve both upon this estimate and upon the theory by investigating the random link TSP using replica techniques often employed in spin glass problems. This approach allowed them to obtain, via a saddle point approximation, many orders of the high-temperature expansion for the internal energy. They then extrapolated down to zero temperature — corresponding to the global TSP optimum — finding $L_{RL} = 1.04 \pm 0.015$. This analysis, like that of Kirkpatrick and Toulouse, was carried out only for the case of constant $\rho(l)$.

Given the difficulties of pushing the replica method further, Mézard and Parisi then tried a different but related approach known as the *cavity* method (MÉZARD & PARISI, 1986*b*). This uses a mean-field approximation which, in the case of spin glasses, gives the same result as the replica method in the thermodynamic limit ($N \to \infty$). Let us sketch what is involved in the cavity method, not least so that we may enumerate clearly the assumptions made.

Both the replica and the cavity approaches involve mapping the TSP onto an m-component spin system, writing down the partition function at temperature T, and then taking the limit $m \to 0$. More explicity, consider N spins $\mathbf{S}_i$, $i = 1, \ldots, N$ (corresponding to the N cities), where each spin $\mathbf{S}_i$ has m components S_i^α, $\alpha = 1, \ldots, m$, and where $\mathbf{S}_i \cdot \mathbf{S}_i = m$ for all i. The partition function is defined, in terms of a parameter ω, as

$$Z = \int d\mu\{\mathbf{S}\} \exp(\omega \sum_{i<j} R_{ij}\, \mathbf{S}_i \cdot \mathbf{S}_j) \tag{III.3}$$

$$= \int d\mu\{\mathbf{S}\} \left[1 + \omega \sum_{i<j} R_{ij} (\mathbf{S}_i \cdot \mathbf{S}_j) + \frac{\omega^2}{2!} \sum_{\substack{i<j \\ k<l}} R_{ij}\, R_{kl} (\mathbf{S}_i \cdot \mathbf{S}_j)(\mathbf{S}_k \cdot \mathbf{S}_l) + \cdots \right] \tag{III.4}$$

[1] Here we work in units where the line segment is taken to have unit length, and in order to match the normalized 1-D Euclidean distribution, we let $\rho(l) = 2$ on $(0, 1/2)$. The 1-D Euclidean value, for comparison, would thus be $L_E = 1$. Kirkpatrick and Toulouse, among others, choose instead $\rho(l) = 1$ on $(0, 1)$, contributing an additional factor of 2 in L_{RL} which we omit when quoting their results.

where $\int d\mu\{\mathbf{S}\}$ denotes an integral over all possible spin values (this simply corresponds to the surface of an m-dimensional sphere), and R_{ij} is related to the length l_{ij} between city i and city j as $R_{ij} \equiv e^{-N^{1/d} l_{ij}/T}$. Now we employ a classic diagrammatic argument: let each spin product $(\mathbf{S}_p \cdot \mathbf{S}_q)$ appearing in the series be represented by an edge in a graph. The first-order terms (ω) will consist of one-edge diagrams, the second-order terms (ω^2) will consist of two-edge diagrams, and so on (see Figure III–1). What then happens when we integrate over all spin configurations? If there is a spin $\mathbf{S}_p$ which occurs only once in a given diagram, *i.e.*, it is an endpoint, the spherical symmetry of $\mathbf{S}_p$ will cause the whole expression to vanish. We will therefore be left only with "closed" diagrams, where there is at least one loop. It may furthermore be shown that in performing the integration, any one of these closed diagrams will contribute a factor m for every loop present in the diagram (De Gennes, 1972; Orland, 1985). If we then consider $(Z-1)/m$ and take the limit $m \to 0$, it is clear that only diagrams with a single loop will remain. Furthermore, since any closed diagram with more than N links must necessarily contain more than one loop, only diagrams up to order ω^N will remain. Finally, take the limit $\omega \to \infty$. The term that will then dominate in (III.4) is the order ω^N term which, being a single loop diagram, represents precisely a closed tour passing through all N sites. We may write it without the combinatorial factor $N!$ by expressing it as a sum over ordered pairs in the tour, and we thus find:

$$\lim_{\substack{m\to 0 \\ \omega\to\infty}} \frac{Z-1}{m\omega^N} = \sum_{\substack{N\text{-link single loops} \\ (i_1,i_2,\ldots,i_n)}} R_{i_1 i_2} R_{i_2 i_3} \cdots R_{i_{N-1} i_N} R_{i_1 i_N} \tag{III.5}$$

$$= \sum_{N\text{-city tours}} e^{-N^{1/d} L/T} \tag{III.6}$$

where L is the total tour length. We thus obtain exactly the partition function for the traveling salesman problem, with the correct canonical ensemble Boltzmann weights, using the tour length as the energy to be minimized (up to a factor $N^{1/d}$, necessary for the energy to be extensive).

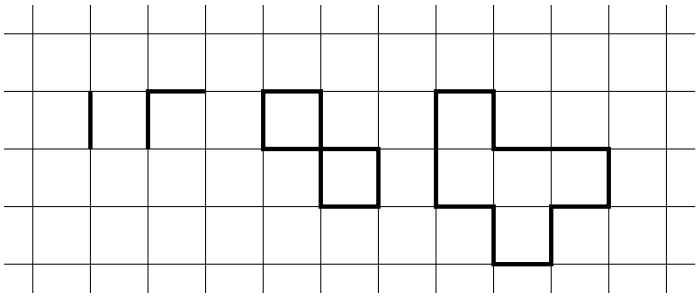

Figure III–1: A graphical representation of a 1-edge open diagram; a 2-edge open diagram; an 8-edge closed diagram with two loops; and a 12-edge closed diagram with one loop.

The cavity method now consists of adding an additional $(N+1)$th spin to the system in the presence of an infinitesimally small external magnetic field oriented along component 1, leading to a spontaneous magnetization $\langle S_i^1 \rangle$. The magnetization $\langle S_{N+1}^1 \rangle$ is then expressed in terms of what the other $\langle S_i^1 \rangle$'s would be in the absence of site $N+1$, hence the notion of a "cavity". In order to do this, two important assumptions are made. First of all, we assume that there is one single equilibrium state. That is to say, at all temperatures a unique thermodynamic limit exists. A similar assumption of ergodicity is made to obtain the replica solution given earlier. The assumption is known as replica symmetry: if a large number of "copies" of the system are allowed to develop, they all reach the same equilibrium state. The notion of replica symmetry is central to the study of disordered systems. Replica symmetry is known to be broken in spin glasses, though there is no reason why this should necessarily be so in other NP-hard problems. Showing that the cavity solution correctly predicts macroscopic quantities for the random link TSP suggests that indeed the TSP exhibits replica symmetry.

The second important assumption on which the cavity method is based involves neglecting the correlations between spins $\mathbf{S}_i, i = 1, \ldots, N$, when calculating $\langle S_{N+1}^1 \rangle$. The assumption is that at large N, the contributions due to these correlations are $O(1/N)$ or less, so that when $N \to \infty$ this sort of mean-field approximation is justified. Proceeding in this way, one obtains from the partition function a recursion relation for $\langle S_{N+1}^1 \rangle$ in terms of the other $\langle S_i^1 \rangle$'s:

$$\langle S_{N+1}^1 \rangle = \frac{\sum_{i=1}^{N} R_{N+1,i} \langle S_i^1 \rangle}{\sum_{1 \le i < j \le N} R_{N+1,i} \langle S_i^1 \rangle R_{N+1,j} \langle S_j^1 \rangle}. \tag{III.7}$$

Furthermore, it may be seen from (III.4) that the diagrams contributing to the thermal average $\langle \mathbf{S}_p \cdot \mathbf{S}_q \rangle$ are those that have link p—q occupied; $\langle \mathbf{S}_p \cdot \mathbf{S}_q \rangle$ is then the occupation number n_{pq} of the link. Since the spontaneous magnetization is along component 1, n_{pq} is simply $\langle S_p^1 S_q^1 \rangle$. From the partition function, a recursion relation for $n_{N+1,i}$ in terms of the $\langle S_i^1 \rangle$'s may be obtained much as in (III.7):

$$n_{N+1,i} = R_{N+1,i} \langle S_{N+1}^1 \rangle \langle S_i^1 \rangle \frac{\sum_{j \neq i} R_{N+1,j} \langle S_j^1 \rangle}{\sum_{j=i}^{N} R_{N+1,j} \langle S_j^1 \rangle}. \tag{III.8}$$

Now consider the effect of our second (mean-field) assumption over the ensemble of instances (distribution over the disorder). As far as (III.7) is concerned, we may treat the magnetizations $\langle S_i^1 \rangle$ as independent identically distributed random variables. Requiring $\langle S_{N+1}^1 \rangle$ to have the same distribution as the others imposes, for a given link length distribution $\rho(l)$, a unique self-consistent probability distribution of the magnetizations. Via (III.8), then, the distribution $\mathcal{P}_d(l)$ of link lengths l *in the optimal tour* (at $N \to \infty$) may be found. KRAUTH & MÉZARD (1989) carried out this calculation in the $T \to 0$ limit, for $\rho(l)$ corresponding to that of the d-dimensional Euclidean case, namely

$$\rho(l) = \frac{d\,\pi^{d/2}}{\Gamma(d/2+1)}\, l^{d-1}. \tag{III.9}$$

They found

$$\mathcal{P}_d(l) = N^{-1/d}\pi^{d/2}\,\frac{\Gamma(d/2+1)}{\Gamma(d+1)}\,\frac{l^{d-1}}{2\Gamma(d)} \tag{III.10}$$

$$\times\left(-\frac{\partial}{\partial l}\right)\int_{-\infty}^{+\infty}\left[1+H_d(x)\right]e^{-H_d(x)}\left[1+H_d(l-x))\right]e^{-H_d(l-x)}\,dx, \tag{III.11}$$

where $H_d(x)$ is the solution to the integral equation

$$H_d(x) = \pi^{d/2}\,\frac{\Gamma(d/2+1)}{\Gamma(d+1)}\int_{-x}^{+\infty}\frac{(x+y)^{d-1}}{\Gamma(d)}\,\left[1+H_d(y)\right]\,e^{-H_d(y)}\,dy. \tag{III.12}$$

From $\mathcal{P}_d(l)$, one may obtain the mean link length in the tour, and thus the total length of the tour:

$$L_{RL}(N,d) = N\int_0^{+\infty} l\,\mathcal{P}_d(l)\,dl \tag{III.13}$$

$$= N^{1-1/d}\,\frac{d}{2}\int_{-\infty}^{+\infty} H_d(x)\left[1+H_d(x)\right]e^{-H_d(x)}\,dx \tag{III.14}$$

for large N. At $d=1$, Krauth and Mézard solved these equations numerically, obtaining $L_{RL} = 1.0208$. It is difficult to compare this with Kirkpatrick's value of 1.045 from direct simulations (as no error estimate exists for the latter quantity), however Johnson, McGeoch & Rothberg (1996) have recently obtained the numerical result $L_{RL}(N = 10{,}000) = 1.0211 \pm 0.0003$, giving strong credence to the cavity value. Krauth and Mézard also performed a numerical study of $\mathcal{P}_1(l)$. They found the cavity predictions to be in good agreement with what they found in their own direct simulations. Further numerical evidence supporting the assumption of replica symmetry was found in an analysis, by Sourlas (1986), of the low temperature statistical mechanics of the system. Thus, for the l_{ij} distribution at $d=1$, there is good reason to believe that the cavity assumptions are valid and that the resulting predictions are exact at large N.

The cavity solution was extended to higher dimensions by the present authors (Percus & Martin, 1996; Cerf, Boutet de Monvel, Bohigas, Martin & Percus, 1997), and a large d power series solution was given for the asymptotic value $\beta_{RL}(d) \equiv \lim_{N\to\infty} L_{RL}(N,d)/N^{1-1/d}$:

$$\beta_{RL}(d) = \sqrt{\frac{d}{2\pi e}}\,(\pi d)^{1/2d}\,\left[1+\frac{2-\ln 2-2\gamma}{d}+O\left(\frac{1}{d^2}\right)\right], \tag{III.15}$$

where γ represents Euler's constant ($\gamma \approx 0.57722$). It would be nice, then, to have evidence that the cavity method is exact for *all* d and not just at $d=1$. While certain arguments have been advanced in favor of this claim (Mézard, Parisi & Virasoro, 1987; Sourlas, 1986), they have yet to be backed up by numerical evidence — as they have been, for instance, in a related combinatorial optimization problem known as the matching problem (Brunetti, Krauth, Mézard & Parisi, 1991; Boutet de Monvel, 1996). We now turn to this task, considering first the $d=2$ case, and then a "renormalized" random link model enabling us to verify numerically the $O(1/d)$ coefficient in (III.15).

– 3 – Numerical analysis: $d = 2$ case

We have implicitly been making the assumption so far, via our notation, that as $N \to \infty$ the random variable $L_{RL}(N,d)/N^{1-1/d}$ approaches a unique value $\beta_{RL}(d)$ with probability 1. This is a property known as self-averaging. The analogous property has been shown for the Euclidean TSP at all dimensions (Beardwood, Halton & Hammersley, 1959). For the random link TSP, however, the only case where a proof of self-averaging is known is in the $d \to \infty$ limit, where a converging upper and lower bound give in fact the *exact* result (Vannimenus & Mézard, 1984):

$$\beta_{RL}(d) = \sqrt{\frac{d}{2\pi e}}\,(\pi d)^{1/2d}\left[1 + O\left(\frac{1}{d}\right)\right]. \tag{III.16}$$

Note, incidentally, that this already shows that the cavity solution (III.15) is correct in the infinite dimensional limit.

For finite d, however, it has not been shown analytically that $\beta_{RL}(d)$ even exists. To some extent, the difficulty in proving this can be traced to the non-satisfaction of the triangle inequality. The reader acquainted with the proof by Beardwood, *et al.*, may see that the ideas used there are not applicable to the random link case; for instance, combining good subtours using

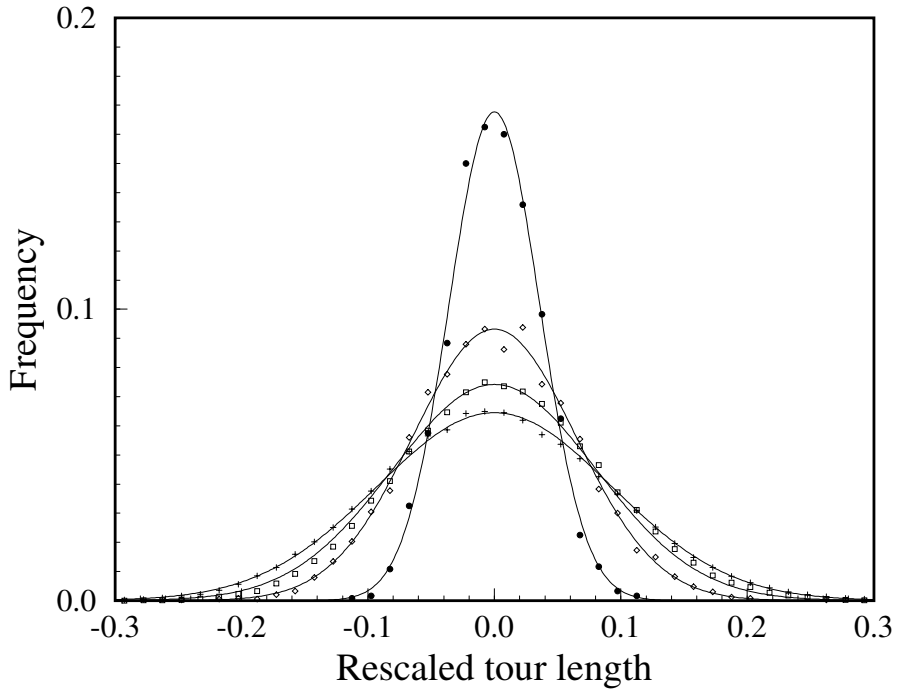

Figure III–2: Distribution of 2-D random link rescaled tour length $(L_{RL} - \langle L_{RL}\rangle)/\sqrt{N}$ for increasing values of N. Plus signs show $N = 12$ (100,000 instances used), squares show $N = 17$ (100,000 instances used), diamonds show $N = 30$ (4,000 instances used), and dots show $N = 100$ (1,200 instances used). Solid lines represent Gaussian fits for each value of N plotted.

simple insertions will not lead to near-optimal global tours, making the problem particularly challenging. Let us therefore examine the distribution of $d = 2$ optimum tour lengths using numerical simulations, in order to give empirical support for the assertion that the $N \to \infty$ limit is well-defined.

In Figure III–2, we see that the $L_{RL}(N, 2)/\sqrt{N}$ distribution indeed becomes increasingly sharply peaked for increasing N. Furthermore the variance of $L_{RL}(N, 2)$ remains relatively constant in N (see Table III–1), indicating that the width σ for the distribution shown in the figure decreases as $1/\sqrt{N}$, strongly suggesting a Gaussian distribution. Similar results were found in an earlier study of ours concerning the Euclidean TSP (CERF, BOUTET DE MONVEL, BOHIGAS, MARTIN & PERCUS, 1997) (albeit in that case with σ approximately half of its random link value). This is precisely the sort of behavior one would expect were the central limit theorem to be applicable.

Table III–1: Variance of the non-rescaled optimum tour length $L_{RL}(N, 2)$ with increasing N.

N	σ^2	# instances used
12	0.3200	100,000
17	0.3578	100,000
30	0.3492	4,000
100	0.3490	1,200

It is worth noting that self-averaging in a different quantity, the Parisi overlap q (MÉZARD, PARISI, SOURLAS, TOULOUSE & VIRASORO, 1984; MÉZARD, PARISI & VIRASORO, 1987), would in fact be a direct indication of replica symmetry. The overlap q_{ab} between two states (tours) a and b may be defined in the TSP, following Kirkpatrick and Toulouse, as the fraction of links that are common to tours a and b. If we then consider suboptimal tours produced by a finite-temperature algorithm, and measure the overlaps between these tours, replica symmetry requires that the overlap distribution become more and more sharply peaked as $N \to \infty$. Results by SOURLAS (1986) confirm that this is so for $d = 1$, at temperatures going down to $T = 0.85$. Unfortunately, such simulations at lower temperatures do not appear feasible: Sourlas' results suggest that fluctuations in q become uncontrollable as $T \to 0$. As we are interested primarily in the $T = 0$ case, we do not consider q in the present analysis, resorting instead to direct tests of the cavity predictions.

The algorithmic procedures we use for simulations are identical to those we used in our Euclidean study; for details, the interested reader is referred to that article. Let us, however, note here that our optimization procedure involves using the LK and CLO local search heuristic algorithms (LIN & KERNIGHAN, 1973; MARTIN & OTTO, 1996) where for each instance of the ensemble we run the heuristic over multiple random starts. LK is used for smaller values of N ($N \leq 17$) and CLO, a more sophisticated method combining LK optimization with random jumps, for larger values of N ($N = 30$ and $N = 100$). Practitioners may note at this point that given the problems we have mentioned in obtaining near-optimal tours from near-optimal subtours in

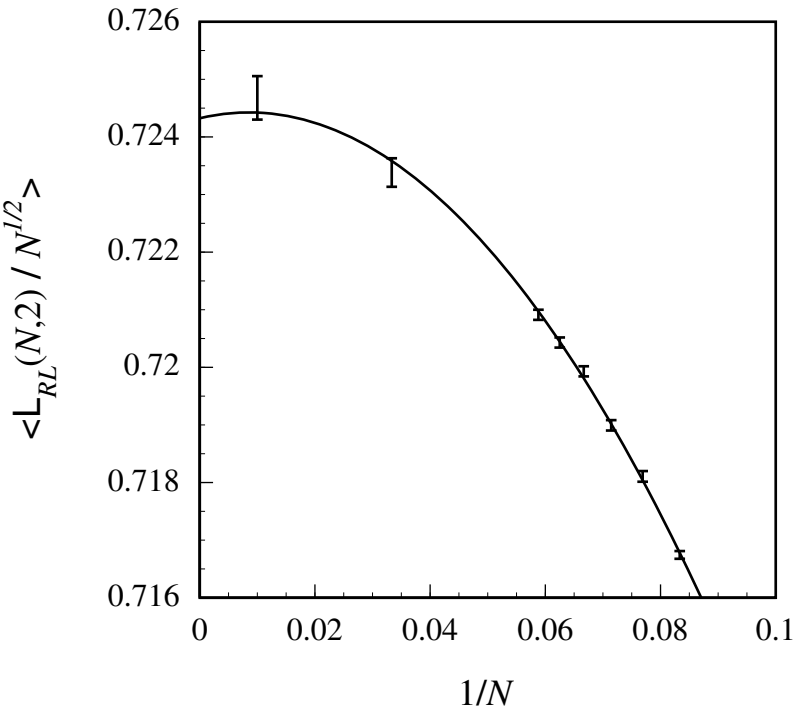

Figure III–3: Finite size scaling of $d = 2$ optimum. Best fit ($\chi^2 = 4.46$) is given by: $\langle L_{RL}(N,2)\rangle/N^{1/2} = 0.7243(1 + 0.0322/N - 1.886/N^2)$. Error bars show one standard deviation (statistical error).

the random link TSP, local search methods would be expected to perform very poorly, leading to relative excess lengths which diverge as $N \to \infty$ and thus next to useless for a large N analysis. In the case of the LK heuristic, interestingly enough, the divergence appears to be no worse than logarithmic in N (JOHNSON & MCGEOCH, 1997), presumably because of the "variable depth" search that LK conducts. There is, nevertheless, a certain probability that even over the course of multiple random starts, our heuristics will not find the true optimum of an instance. We estimate the associated systematic bias using a number of test instances, and adjust the number of random starts to keep this bias at least an order of magnitude below other sources of error discussed below. (At its maximum — occurring in the $N = 100$ case — the systematic bias is estimated as under 1 part in 20,000.)

Following this numerical method, let us now consider the large N limit of $L_{RL}(N,2)$. In the Euclidean case, it has been observed that the finite size scaling law can be written in terms of a power series in $1/N$. We expect this same behavior in the random link case, namely that the ensemble average $\langle L_{RL}(N,2)\rangle$ satisfies

$$\langle L_{RL}(N,d)\rangle = \beta_{RL}(d)\, N^{1-1/d} \left[1 + \frac{A(d)}{N} + \cdots\right]. \tag{III.17}$$

In order to obtain $\langle L_{RL}(N,2)\rangle$ at a given value of N from simulations, we average over a large

number of instances to reduce the statistical error arising from instance-to-instance fluctuations. Figure III–3 shows the results of this, with accompanying error bars, fitted to the expected finite size scaling law. The fit is a good one ($\chi^2 = 4.46$ for 5 degrees of freedom) and gives an extrapolated value of $\beta_{RL}(2) = 0.7243 \pm 0.0004$. This is in superb agreement with the cavity result of 0.7251. The relative discrepancy between the cavity prediction and our numerical estimate is approximately 0.1%, which in consistent with our statistical error bars; by comparison, the error in the replica symmetric solution to the Sherrington-Kirkpatrick spin glass ground state energy is estimated to be of the order of 5% (Mézard, Parisi & Virasoro, 1987, pp. 13–14).

Another quantity that was considered in the $d = 1$ numerical study of Krauth & Mézard (1989) is the optimal tour link length distribution $\mathcal{P}_d(l)$ given in (III.10). Let us consider $\mathcal{P}_2(l)$, and following their example, let us look specifically at the integrated distribution $I_d(l) \equiv \int_0^l \mathcal{P}_d(\tilde{l})\, d\tilde{l}$. The cavity result for $I_d(l)$ can, like $\beta_{RL}(d)$, be computed numerically to arbitrary precision. In Figure III–4 we compare this with the results of direct simulations, at increasing values of N. The improving agreement for increasing N (already within 2% at $N = 100$) strongly suggests that the cavity solution gives the exact $N \to \infty$ result.

Finally, it may be of interest to consider one further quantity in the $d = 2$ random link simulations: the frequencies of "neighborhoods" used in the optimal tour, that is, the proportion of links connecting nearest neighbors, 2nd-nearest neighbors, etc. While there is no cavity

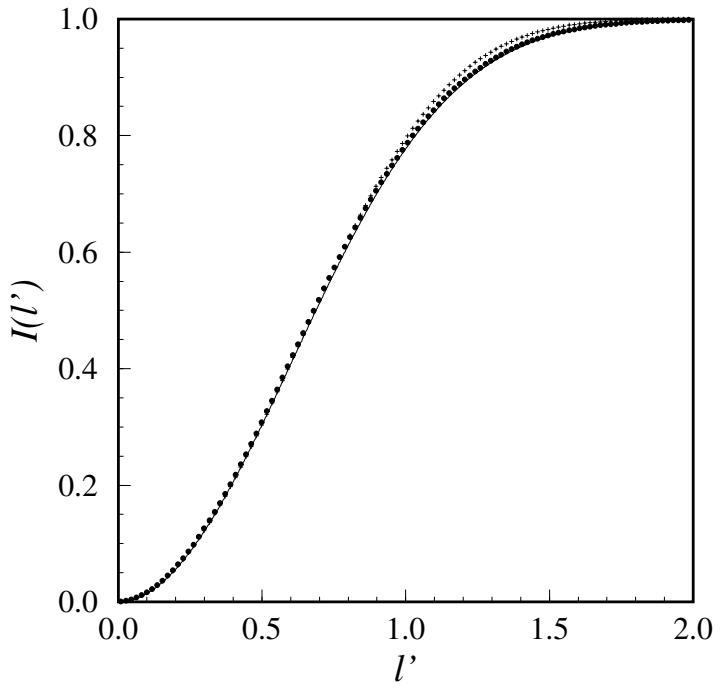

Figure III–4: Integrated probability distribution of link lengths in the optimal tour, for $d = 2$, using rescaled length $l' = l\sqrt{N}$. Plus signs represent $N = 12$ simulation results, dots represent $N = 100$ simulation results, and solid line represents cavity prediction.

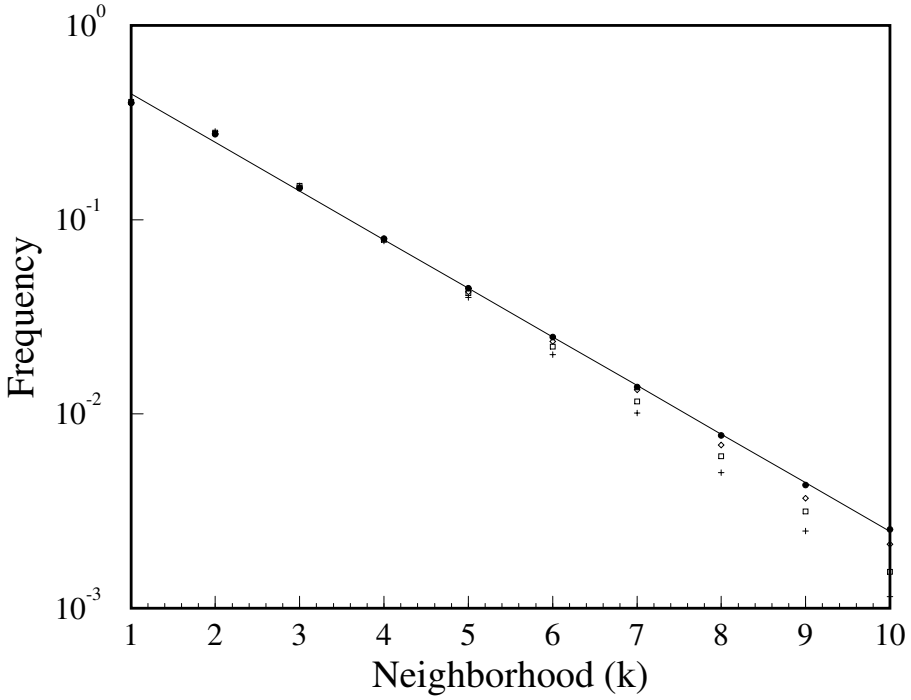

Figure III–5: Frequencies with which kth-nearest neighbors are used in optimal 2-D random link tours. Plus signs show values for $N = 12$, squares for $N = 17$, diamonds for $N = 30$, and dots for $N = 100$. Best exponential fit (appearing linear on log plot) is shown for $N = 100$ data.

prediction for this quantity, Sourlas (1986) has noted that in practice in the $d = 1$ case, the frequency falls off rapidly with increasing neighborhood — suggesting that optimization heuristics could be improved by preferentially choosing links between very near neighbors. Results of our simulations, shown in Figure III–5, suggest that at $d = 2$ the decrease is in fact very close to exponential. We have no theoretical explanation for this behavior at present, and consider it a fascinating open question.

— 4 — Numerical analysis: renormalized model

Let us now consider a different sort of random link tsp, proposed in Cerf, Boutet de Monvel, Bohigas, Martin & Percus (1997), allowing us to test numerically the $1/d$ coefficient in the cavity result (III.15). Define $\langle D_1(N,d)\rangle$ to be the distance between a city and its nearest neighbor, averaged over all cities in the instance and over all instances in the ensemble.[2] For large d, it may be shown that

$$\lim_{N\to\infty} N^{1/d}\langle D_1(N,d)\rangle = \sqrt{\frac{d}{2\pi e}}\,(\pi d)^{1/2d}\left[1 - \frac{\gamma}{d} + O\left(\frac{1}{d^2}\right)\right] \tag{III.18}$$

[2]Note that $\langle D_1(N,d)\rangle$ itself does not involve the notion of optimal tours, or tours of any sort for that matter.

where γ is Euler's constant. It is not surprising that this quantity is reminiscent of (III.16), since $N^{1/d}\langle D_1(N,d)\rangle$ represents precisely a lower bound on $\beta_{RL}(d)$.

In the renormalized random link model, we define a new link length $x_{ij} \equiv d[l_{ij} - \langle D_1(N,d)\rangle]/\langle D_1(N,d)\rangle$, where the l_{ij} have the usual distribution corresponding to d dimensions. The x_{ij} are "lengths" only in the loosest sense, as they can be both positive and negative. The optimal tour in the x_{ij} model will, however, follow the same "path" as the optimal tour in the associated l_{ij} model since the transformation is linear. Its length $L_x(N,d)$ will simply be given in terms of $L_l(N,d)$ by:

$$L_x(N,d) = d\,\frac{L_l(N,d) - N\langle D_1(N,d)\rangle}{\langle D_1(N,d)\rangle}, \text{ so} \tag{III.19}$$

$$L_l(N,d) = N\langle D_1(N,d)\rangle\left[1 + \frac{L_x(N,d)}{dN}\right]. \tag{III.20}$$

By definition, then,

$$\beta_{RL}(d) = \lim_{N\to\infty} N^{1/d}\langle D_1(N,d)\rangle, \text{ or at large } d, \tag{III.21}$$

$$= \sqrt{\frac{d}{2\pi e}}\,(\pi d)^{1/2d}\left[1 - \frac{\gamma}{d} + O\left(\frac{1}{d^2}\right)\right]\lim_{N\to\infty}\left[1 + \frac{L_x(N,d)}{dN}\right]. \tag{III.22}$$

If $\beta_{RL}(d)$ is to be well-defined, then there must exist a value $\mu(d)$ such that $\lim_{N\to\infty} L_x(N,d)/N = \mu(d)$.

What will be the distribution of lengths $\rho(x)$ corresponding to $\rho(l)$? From the definition of $\rho(l)$ and that of x,

$$\rho(x) = \frac{d\,\pi^{d/2}\,l^{d-1}}{\Gamma(d/2+1)}\,\frac{\langle D_1(N,d)\rangle}{d}, \text{ and substituting for } l, \tag{III.23}$$

$$= \frac{\pi^{d/2}}{\Gamma(d/2+1)}\left(1 + \frac{x}{d}\right)^{d-1}\langle D_1(N,d)\rangle^d. \tag{III.24}$$

In the limit $N \to \infty$, this gives

$$\rho(x) \sim \frac{\pi^{d/2}}{\Gamma(d/2+1)}\left(1 + \frac{x}{d}\right)^{d-1} N^{-1}\left(\frac{d}{2\pi e}\right)^{d/2}\sqrt{\pi d}\left[1 - \frac{\gamma}{d} + \cdots\right]^d \tag{III.25}$$

$$\sim N^{-1}\left(1 - \frac{\gamma}{d}\right)^d\left(1 + \frac{x}{d}\right)^{d-1}\left[1 + O\left(\frac{1}{d}\right)\right] \text{ by Stirling's formula} \tag{III.26}$$

$$\sim N^{-1}e^{x-\gamma}\left[1 + O\left(\frac{1}{d}\right)\right] \tag{III.27}$$

At large d, we see that $\rho(x)$ is to leading order independent of d; the same must then be true for $L_x(N,d)$, and hence for $\mu(d)$, which we now write as $\mu_0[1 + O(1/d)]$. From (III.22), we obtain

$$\beta_{RL}(d) = \sqrt{\frac{d}{2\pi e}}\,(\pi d)^{1/2d}\left[1 + \frac{\mu_0 - \gamma}{d} + O\left(\frac{1}{d^2}\right)\right]. \tag{III.28}$$

Finally, we may perfectly well define $\rho(x)$ in the $d \to \infty$ limit, in which case (III.27) gives $\rho(x) = N^{-1}\exp(x-\gamma)$. Doing so will give us a renormalized model with $\lim_{N\to\infty} L_x(N)/N = \mu_0$.

By performing direct simulations on a random link model with this exponential distribution, we may find the value of μ_0 numerically, and thus from (III.28) the $1/d$ coefficient for the d-dimensional (non-renormalized) case.

Figures III–6 and III–7 show the numerical results for the renormalized model. In Figure III–6, we see that just as in the $d = 2$ case, the distribution of the optimal tour length becomes sharply peaked at large N and the asymptotic limit μ_0 is well-defined. Via (III.22), this provides very good reason for believing that $\beta_{RL}(d)$ is well-defined *for all* d, and self-averaging holds for the random link TSP in general. In Figure III–7, we show the finite size scaling of $\langle L_x(N)/N \rangle$. The fit is again quite satisfactory (with χ^2=5.23 for 5 degrees of freedom), giving the asymptotic result $\mu_0 = 0.7300 \pm 0.0010$. The simulated value for the $1/d$ coefficient in $\beta_{RL}(d)$ is then $\mu_0 - \gamma = 0.1528 \pm 0.0010$, in excellent agreement (error under 0.3%) with the cavity value $2 - \ln 2 - 2\gamma \approx 0.1524$ given in (III.15).

Also as in the $d = 2$ case, let us briefly consider the frequencies of kth-nearest neighbors used in optimal tours. This is given in Figure III–8. Even though the exponential fit is not as excellent as in the $d = 2$ case, it is still striking here. We have seen that the renormalized random link model is analogous to the $d \to \infty$ limit of the standard random model, as far as tour paths are concerned. These kth-neighbor frequency results then give empirical evidence

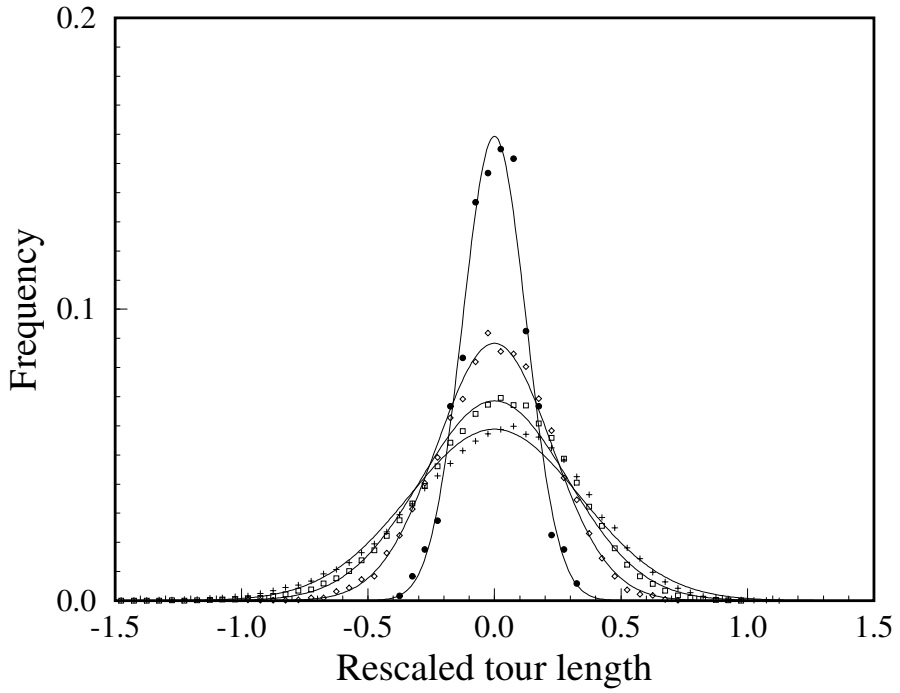

Figure III–6: Distribution of renormalized random link rescaled tour length $(L_x - \langle L_x \rangle)/N$ for increasing values of N. Plus signs show $N = 12$ (100,000 instances used), squares show $N = 17$ (100,000 instances used), diamonds show $N = 30$ (4,000 instances used), and dots show $N = 100$ (1,200 instances used). Solid lines represent Gaussian fits for each value of N plotted.

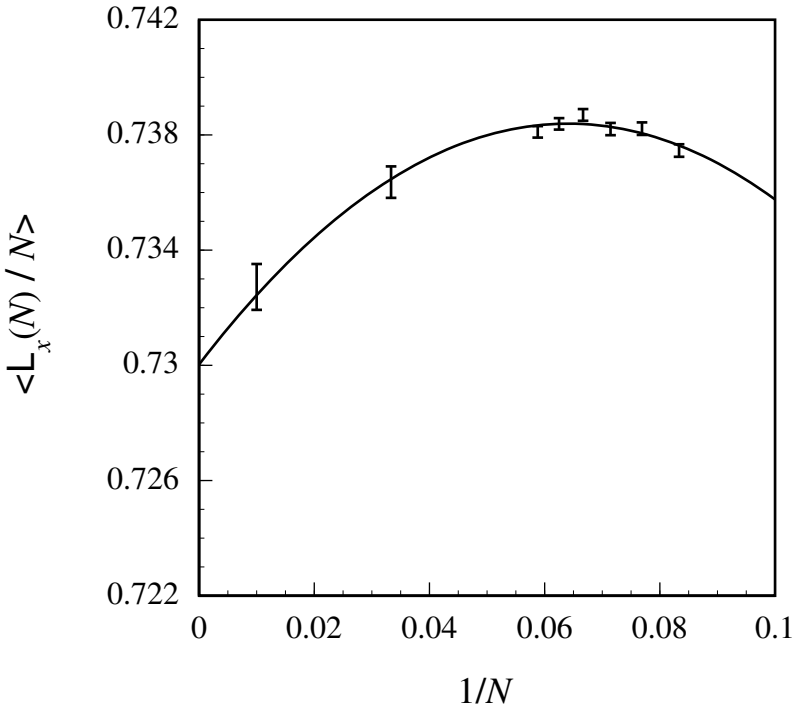

Figure III–7: Finite size scaling of "renormalized model" optimum. Best fit ($\chi^2 = 5.23$) is given by: $\langle L_x(N)\rangle/N = 0.7300(1 + 0.3575/N - 2.791/N^2)$. Error bars show one standard deviation (statistical error).

that even at infinite dimension, the "typical" k used remains bounded.

— 5 — Conclusion

The random link TSP has interested theoreticians primarily because of its analytical tractability, allowing presumably exact results that are not possible in the more traditional Euclidean TSP. Other than in the $d = 1$ case, however, it has attracted little attention from the angle of numerics. In this paper we have provided a numerical study of the random link TSP that was lacking up to this point, and which addresses some long-unanswered questions. Through simulations, we have tested the validity of the theoretical cavity method predictions. While in other disordered systems, such as spin glasses, the replica symmetric solution gives values of macroscopic quantities that are inexact (typically by at least 5%), in the random link TSP it shows all signs of being exact. We have studied various macroscopic quantities at $d = 2$ and found that the numerical results confirm the cavity predictions to within 0.1%. Furthermore, we have confirmed, by way of a renormalized random link model, that the analytical cavity solution gives a large d expansion for the optimal tour length whose $1/d$ coefficient is correct to within well

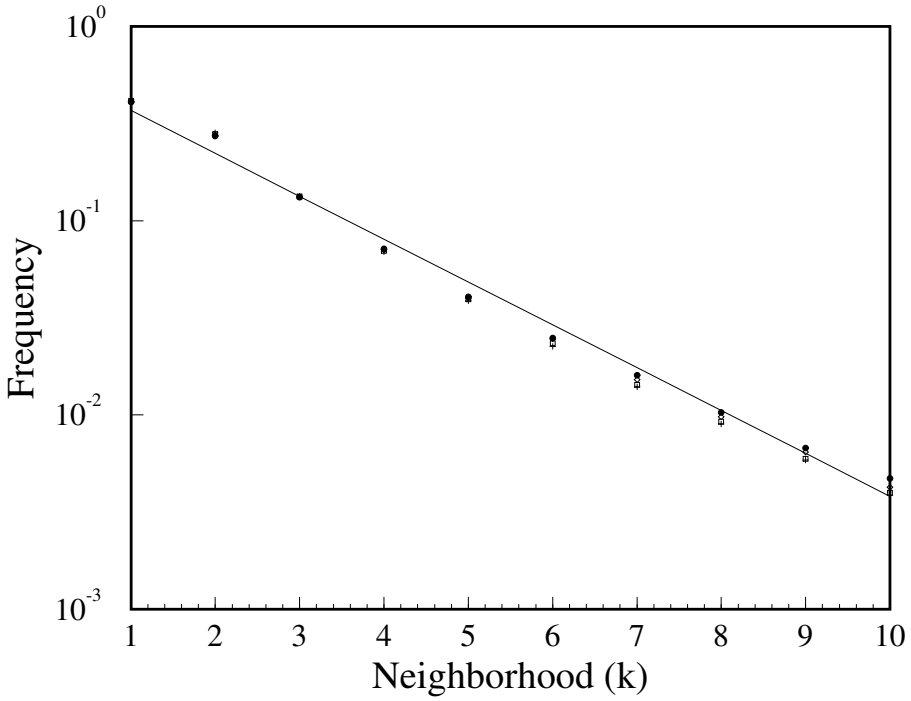

Figure III–8: Frequencies with which kth-nearest neighbors are used in optimal renormalized random link tours. Plus signs show values for $N = 12$, squares for $N = 17$, diamonds for $N = 30$, and dots for $N = 100$. Best exponential fit (appearing linear on log plot) is shown for $N = 100$ data.

under 1%. The excellent agreement found at $d = 1$ (Krauth & Mézard, 1989; Sourlas, 1986), $d = 2$, and to $O(1/d)$ at large d, then suggest strongly that the cavity predictions are exact. This, finally, provides indirect evidence that the assumption of replica symmetry — on which the cavity approach is based — is indeed justified.

Bibliography

Beardwood J., Halton J.H. & Hammersley J.M. (1959): "The Shortest Path through Many Points", *Proc. Cambridge Philos. Soc.*, **55**, pp. 299–327.

Boutet de Monvel J.H. (1996): "Physique Statistique et Modèles à Liens Aléatoires", Ph.D. thesis, Université Paris-Sud, Orsay, France.

Brunetti R., Krauth W., Mézard M. & Parisi G. (1991): "Extensive Numerical Solutions of Weighted Matchings: Total Length and Distribution of Links in the Optimal Solution", *Europhys. Lett.*, **14**(4), pp. 295–301.

Cerf N.J., Boutet de Monvel J., Bohigas O., Martin O.C. & Percus A.G. (1997): "The Random Link Approximation for the Euclidean Traveling Salesman Problem", *J. Phys. I France*, **7**(1), pp. 117–136.

De Gennes P.G. (1972): "Exponents for the Excluded Volume Problem as Derived by the Wilson Method", *Phys. Lett. A*, **38**(5), pp. 339–340.

Johnson D.S. & McGeoch L.A. (1997): "The Traveling Salesman Problem: A Case Study in Local Optimization", in *Local Search in Combinatorial Optimization*, ed. by Aarts E.H.L. & Lenstra J.K. New York: John Wiley and Sons, to appear.

Johnson D.S., McGeoch L.A. & Rothberg E.E. (1996): "Asymptotic Experimental Analysis for the Held-Karp Traveling Salesman Bound", in *7th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 341–350, Atlanta, GA.

Kirkpatrick S. & Toulouse G. (1985): "Configuration Space Analysis of Travelling Salesman Problem", *J. Phys. France*, **46**, pp. 1277–1292.

Krauth W. & Mézard M. (1989): "The Cavity Method and the Travelling-Salesman Problem", *Europhys. Lett.*, **8**, pp. 213–218.

Lin S. & Kernighan B. (1973): "An Effective Heuristic Algorithm for the Traveling Salesman Problem", *Operations Res.*, **21**, pp. 498–516.

Martin O.C. & Otto S.W. (1996): "Combining Simulated Annealing with Local Search Heuristics", *Ann. Operations Res.*, **63**, pp. 57–75.

Mézard M. & Parisi G. (1986*a*): "A Replica Analysis of the Travelling Salesman Problem", *J. Phys. France*, **47**, pp. 1285–1296.

——— (1986*b*): "Mean-Field Equations for the Matching and the Travelling Salesman Problems", *Europhys. Lett.*, **2**, pp. 913–918.

Mézard M., Parisi G., Sourlas N., Toulouse G. & Virasoro M. (1984): "Replica Symmetry Breaking and the Nature of the Spin Glass Phase", *J. Phys. France*, **45**, pp. 843–854.

Mézard M., Parisi G. & Virasoro M.A. (eds.) (1987): *Spin Glass Theory and Beyond*. Singapore: World Scientific.

Orland H. (1985): "Mean-field Theory for Optimization Problems", *J. Phys. Lett. France*, **46**, pp. L763–L770.

Percus A.G. & Martin O.C. (1996): "Finite Size and Dimensional Dependence in the Euclidean Traveling Salesman Problem", *Phys. Rev. Lett.*, **76**, pp. 1188–1191.

Sherrington D. & Kirkpatrick S. (1975): "Solvable Model of a Spin-Glass", *Phys. Rev. Lett.*, **35**(26), pp. 1792–1796.

Sourlas N. (1986): "Statistical Mechanics and the Travelling Salesman Problem", *Europhys. Lett.*, **2**, pp. 919–923.

VANNIMENUS J. & MÉZARD M. (1984): "On the Statistical Mechanics of Optimization Problems of the Travelling Salesman Type", *J. Phys. Lett. France*, **45**, pp. L1145–L1153.

Chapter IV

Universalities in nearest neighbor distances

Contents

Chapter overview 71
Abstract 73
1. Introduction 75
2. Preliminaries 75
3. Flat Surfaces 77
4. Curved Surfaces 78
5. Regge Calculus 82
6. Conclusions 83
Bibliography 83

– *Chapter overview* –

In this final chapter we turn to a problem that does not involve the TSP as such, but is closely inspired by our TSP work and mirrors much of the language we have employed so far. We have seen in Chapter II how the $N \to \infty$ Euclidean optimal tour length, where N represents the number of cities, could be extrapolated from reasonably small sized instances via an understanding of the finite size scaling behavior. We were thus able to obtain precise numerical estimates for this asymptotic value without undue computational effort.

In order to understand the finite size scaling, we decided to set aside temporarily the notion of optimal tours — or, for that matter, tours of any sort. Instead we simply looked at cities placed randomly and independently, with a uniform distribution, and considered the distances between nearest neighboring cities, second-nearest neighboring cities, and so on. What we found was an unexpected sort of universality. If one considers the ensemble average $\langle D_k(N) \rangle$ of the distances between kth-nearest neighbors, the large N scaling law of this quantity is independent of k. Putting it another way, the N-dependence and the k-dependence of $\langle D_k(N) \rangle$ separate; in d dimensions we are left with, up to corrections exponentially small in N,

$$\langle D_k(N) \rangle = \frac{[\Gamma(d/2+1)]^{1/d}}{\sqrt{\pi}} \frac{\Gamma(k+1/d)}{\Gamma(k)} \frac{\Gamma(N)}{\Gamma(N+1/d)}. \tag{IV.1}$$

We now wish to investigate this universality further. How, first of all, does the property depend on the nature of the physical space used? The universality only applies in flat (*i.e.*, Euclidean) spaces, but leads us to notice another interesting and more subtle property on curved surfaces. Let us look at the large N behavior of $\langle D_k(N)\rangle$ (averaged over the whole surface, if the surface is not homogeneous). We find that $\langle D_k(N)\rangle$ scales as a power of N times a series in $1/N$ whose $O(1/N)$ term, while now explicitly depending on k, has the form of a topological invariant: $(\chi(2k+1)-9)/24$, where χ is the Euler characteristic. To $O(1/N)$, then, it is the topology of the surface and not its detailed shape that plays a role in the finite size scaling behavior.

The $1/N$ series describing $\langle D_k(N)\rangle$ may, for a general manifold, be obtained directly from the area $A(l)$ of a geodesic disc of radius l on the manifold. A geodesic disc about some point $\mathbf{x}$ is simply defined as the locus of points whose distance from $\mathbf{x}$ *along the manifold* is smaller than the radius. For a flat surface, of course, $A(l)=\pi l^2$. For a curved surface, $A(l)$ may be written as a series expansion in l; higher-order terms in this expansion will then correspond to higher-order terms in the $1/N$ series. We must therefore find the series expansion of $A(l)$, whose coefficients involve gradients of the surface's Gaussian curvature K at the point $\mathbf{x}$ about which the expansion is being performed. Unfortunately, we know of no general expression for these coefficients. Calculating them is a somewhat laborious procedure, which we have carried out up to the $O(l^8)$ term in $A(l)$:

$$\begin{aligned} A(l) &= \pi l^2 - \frac{\pi l^4}{12}K + \frac{\pi l^6}{720}\big(2K^2 - 3\nabla^2 K\big) \\ &\quad -\frac{\pi l^8}{161280}\big(8K^3 - 3[10(\nabla K)^2 + 14K\nabla^2 K - 5\nabla^4 K]\big) + O(l^{10}). \end{aligned} \tag{IV.2}$$

From this, terms through $O(1/N^3)$ can be found in the $1/N$ series for $\langle D_k(N)\rangle$. We discuss these higher-order terms, although it appears that no further universalities exist beyond the topological invariant at $O(1/N)$. We then turn briefly to higher-dimensional manifolds — where $A(l)$ is far less obvious — and examine the cases in which a topological invariant could exist there. Finally, we provide an interpretation of our results via a Regge calculus approach, where we consider sites distributed on a polyhedral surface rather than a smooth manifold; we find that there too, we recover the topological invariant.

Let us alert the reader to some notational differences in this final chapter that we unfortunately cannot avoid. We find it clearer in the present context to speak of the distance from an arbitrary point $\mathbf{x}$ on the manifold to its kth-nearest site, rather than of the distance between a site and its kth-nearest neighbor. $D_k(N)$ is thus a random variable defined everywhere in (continuous) space, and not only at the positions of one of the N sites. $\langle D_k(N)\rangle$ is its ensemble average, which for inhomogeneous surfaces will vary with $\mathbf{x}$. Our previous notation amounted to restricting $\mathbf{x}$ to be the position of one of the cities, thus reducing the number of possible neighbors by 1. Our results for N cities, in our previous notation, are therefore equivalent to our results for $N-1$ cities, in our new notation. Note, also, that while the terms "cities" and "sites" are interchangeable, they are to be distinguished from the term "point", which is used here in its most precise geometric meaning to represent *any* position $\mathbf{x}$ on the manifold.

Finite size scaling universalities of kth-nearest neighbor distances on closed manifolds

Allon G. Percus and Olivier C. Martin
Division de Physique Théorique *
Institut de Physique Nucléaire
Université Paris-Sud
F-91406 Orsay Cedex, France.

percus@ipno.in2p3.fr
martino@ipno.in2p3.fr

Abstract

Take N sites distributed randomly and uniformly on a closed surface. We express the expected distance $\langle D_k(N)\rangle$ from an arbitrary point on the surface to its kth-nearest neighboring site, in terms of the surface area $A(l)$ of a disc of radius l about that point. We then find two universalities. First, for a flat surface, where $A(l) = \pi l^2$, the finite size scaling series giving corrections to the large N asymptotic behavior of $\langle D_k(N)\rangle$ is independent of k. Second, for a curved surface, the finite size scaling series for the average $\int \langle D_k(N)\rangle \, d\mu$ over the surface is, to $O(1/N)$, a topological invariant. We discuss the case of higher dimensions ($d > 2$), and also interpret our results using Regge calculus.

* Unité de recherche des Universités de Paris XI et Paris VI associée au CNRS

— 1 — Introduction

In many physics and computer science problems, a quantity of interest is the distance between neighboring sites in a space. One frequently wishes to calculate distances to a nearest neighbor, second-nearest neighbor, and more generally, kth-nearest neighbor. Examples in computational geometry and optimization abound, ranging from random packing of spheres to minimum spanning trees. Such problems also arise naturally in physics, both for interacting particle systems such as liquids, and for cellular objects such as foams and random lattices (Itzykson & Drouffe, 1989).

Here we consider the case of N sites placed randomly, with a uniform distribution, on a 2-D surface of fixed area. Let the random variable $D_k(N)$ represent the distance between a given point $\mathbf{x}$ and its kth-nearest neighboring site. Its expectation value $\langle D_k(N)\rangle$ taken over the ensemble of randomly placed sites — and in fact all moments $\langle D_k^\alpha(N)\rangle$ — then exhibit some surprising properties.

When the surface is flat, $\langle D_k(N)\rangle$ at large N is described by its leading asymptotic behavior times a power series in $1/N$, where all orders of the series turn out to be *independent* of k. In the language of statistical physics, the finite size scaling series for $\langle D_k(N)\rangle$ is the same at all k to all orders in $1/N$. Geometrically, this universality is far from obvious. Furthermore, we find that the property is not restricted to two dimensions, and is equally valid for flat spaces of any dimension.

When the surface is curved, this k-independence no longer holds, but one finds on the other hand another universality: the finite size scaling series, averaged over the entire surface, gives a topological invariant to $O(1/N)$. This term thus depends not on the detailed shape of the surface but only on the surface's genus.

In this paper we explore these universalities. First, we express $\langle D_k(N)\rangle$ in terms of the area, $A(l)$, of a disc of radius l on an arbitrary surface. We observe that in the special case where $A(l)$ consists only of a power of l, the $1/N$ series for $\langle D_k(N)\rangle$ exhibits the universality in k, i.e., is independent of k. Second, we give the relation between $A(l)$ and the Gaussian curvature of the surface, and find the leading correction in the $1/N$ power series for $\langle D_k(N)\rangle$, as functions of the curvature. This leads to the topological invariance at $O(1/N)$. We then discuss higher order terms, and the case of higher dimensions. Finally, we show that a Regge calculus approach provides a simple means of obtaining this topological invariance for the case of polyhedral (non-smooth) surfaces.

— 2 — Preliminaries

Take any point $\mathbf{x}$, and consider $P[D_k(N) = l]_{\mathbf{x}}$, the probability density that the point's kth-nearest neighboring site lies at a distance l from it. This is equal to the probability density of having $k-1$ (out of N) sites *within* distance l, one site (out of $N-k+1$) *at* distance l, and the remaining $N-k$ sites *beyond* distance l. Let us choose units so that our surface has area 1. Since sites are distributed uniformly over the surface, the probability of a single site lying within distance l is then simply the area $A(l)_{\mathbf{x}}$ of a disc of radius l about point $\mathbf{x}$ on the surface.

Dropping the argument $\mathbf{x}$ (in order to simplify the notation), we may then write

$$P[D_k(N) = l] = \left(\begin{array}{c} N \\ k-1 \end{array} \right) [A(l)]^{k-1} \times \left(\begin{array}{c} N-k+1 \\ 1 \end{array} \right) \frac{dA(l)}{dl} \times [1 - A(l)]^{N-k},$$

giving the expectation value (first moment)

$$\begin{aligned} \langle D_k(N) \rangle &= \int_0^\infty P[D_k(N) = l]\, l\, dl \\ &= \frac{N!}{(N-k)!\,(k-1)!} \int_0^\infty l\, [A(l)]^{k-1} [1 - A(l)]^{N-k} \frac{dA(l)}{dl}\, dl. \end{aligned}$$

Under the variable transformation $w = A(l)$, this may be written in terms of the inverse function $A^{-1}(w)$ as

$$\langle D_k(N) \rangle = \frac{N!}{(N-k)!\,(k-1)!} \int_0^1 A^{-1}(w)\, w^{k-1} (1-w)^{N-k}\, dw.$$

If $A^{-1}(w)$ admits the power series expansion in w:

$$A^{-1}(w) \simeq w^\gamma \sum_{j=0}^{\infty} c_j w^j, \qquad \text{for some } \gamma \in [0, 1), \tag{IV.3}$$

then

$$\langle D_k(N) \rangle \simeq \frac{N!}{(N-k)!\,(k-1)!} \sum_{j=0}^{\infty} c_j \int_0^1 w^{k+j+\gamma-1} (1-w)^{N-k}\, dw. \tag{IV.4}$$

Recognizing the integral as the Beta function $\mathrm{B}(k+j+\gamma, N-k+1) = (k+j+\gamma-1)!\,(N-k)!/(N+j+\gamma)!$,

$$\langle D_k(N) \rangle \simeq \sum_{j=0}^{\infty} c_j \frac{(k+j+\gamma-1)!}{(k-1)!} \frac{N!}{(N+j+\gamma)!}. \tag{IV.5}$$

Several comments are in order concerning $\langle D_k(N) \rangle$. First of all, although we restrict ourselves to discussing the first moment of $D_k(N)$, we could in fact consider any moment $\langle D_k^\alpha(N) \rangle$ by taking $\left[A^{-1}(w)\right]^\alpha$ instead of $A^{-1}(w)$ in (IV.3). Doing so would alter γ and the c_j's, but would not change our results qualitatively. Second of all, there is no loss of generality in taking our total surface area to be unity; scaling this area by a constant (or even, as might be more intuitive to statistical physicists, by N) would provide only a trivial scaling factor in our results. Third of all, we could imagine that the point $\mathbf{x}$ we consider is itself an $(N+1)$th site. This is simply a question of nomenclature. The problem of finding the expected distance *from* a point *to* its kth-nearest site, for a system of N sites, is therefore equivalent to the problem of finding the expected distance *between* kth-nearest neighbors, for a system of $N+1$ sites.

Let us now define the reduced variable $\langle \tilde{D}_k(N) \rangle$ by dividing out the leading asymptotic (large N) behavior from $\langle D_k(N) \rangle$:

$$\begin{aligned} \langle \tilde{D}_k(N) \rangle &= \langle D_k(N) \rangle \frac{1}{c_0} \frac{(k-1)!}{(k+\gamma-1)!} N^\gamma \\ &\simeq N^\gamma \sum_{j=0}^{\infty} \frac{c_j}{c_0} \frac{(k+j+\gamma-1)!}{(k+\gamma-1)!} \frac{N!}{(N+j+\gamma)!}, \end{aligned} \tag{IV.6}$$

so that $\lim_{N\to\infty} \langle \tilde{D}_k(N) \rangle = 1$ (this is seen from Stirling's law). $\langle \tilde{D}_k(N) \rangle$ then provides the *finite size scaling* behavior for kth-nearest neighbor distances. In what follows, we shall consider the properties of $A^{-1}(w)$, and its consequences on $\langle \tilde{D}_k(N) \rangle$.

— 3 — Flat Surfaces

On a flat surface, if we could neglect edge effects, the area included within distance l would simply be $A(l) = \pi l^2$. In that case, $A^{-1}(w) = \sqrt{w/\pi}$, and so from (IV.3) and (IV.6) we would have

$$\langle \tilde{D}_k(N) \rangle = \frac{\sqrt{N}\, N!}{(N+1/2)!}. \tag{IV.7}$$

The finite size scaling would thus be completely independent of k.

As we are working with a surface of fixed (unit) area, however, we cannot avoid considering edge effects. Let us restrict ourselves to the case where the surface is everywhere locally Euclidean within some minimum neighborhood of radius $l_0 > 0$. (The simplest example of this is a unit square with periodic boundary conditions, for which $l_0 = 1/2$. Obviously, many other constructions are possible.) Any required modifications to the $A^{-1}(w)$ expression in (IV.3) then concern only w greater than $w_0 \equiv A(l_0)$. Correspondingly, (IV.4) remains valid up to remainder terms from the region of integration $w_0 \le w \le 1$. Since the $(1-w)^{N-k-1}$ term in the integral is bounded above by $(1-w_0)^{N-k-1}$ within this region, these remainder terms are exponentially small in N. Equation (IV.7) is therefore still correct *to all orders* in $1/N$, and may be written as the expansion

$$\langle \tilde{D}_k(N) \rangle = 1 - \frac{3}{8N} + O\left(\frac{1}{N^2}\right),$$

where all orders in $1/N$ are independent of k. The finite size scaling law for kth-nearest neighbor distances on a 2-D flat surface without a boundary thus exhibits the universality in k to all orders in $1/N$.

The same holds true for flat manifolds in any dimension d. We assume there is an l_0 such that the volume included within distance $l < l_0$ is simply the volume of a d-dimensional ball:

$$\begin{aligned} A(l) &= \frac{\pi^{d/2}\, l^d}{(d/2)!}, \qquad \text{or} \\ A^{-1}(w) &= \frac{1}{\sqrt{\pi}} \left[w \left(\frac{d}{2}\right)! \right]^{1/d}. \end{aligned}$$

As before, the boundary conditions allow us to write $\langle \tilde{D}_k(N) \rangle$ up to remainder terms that are exponentially small in N, so from (IV.6),

$$\begin{aligned} \langle \tilde{D}_k(N) \rangle &\simeq \frac{N^{1/d}\, N!}{(N+1/d)!} + \cdots \\ &= 1 - \frac{1/d + 1/d^2}{2N} + O\left(\frac{1}{N^2}\right). \end{aligned}$$

Thus for flat spaces without a boundary, of any dimension d, the universality in k holds to all orders in $1/N$.

It may be interesting to consider a slight variation on the problem, giving this universality *exactly* and not only to all orders. Take the case of a spherical surface embedded in 3-D Euclidean space, with the usual measure of area over the sphere, but with a peculiar sort of "distance": rather than the conventional choice of the arc length (geodesic) metric, use the chord length. For a chord of length l originating at a pole of the sphere, the area of the spherical cap spanned by it is simply $A(l) = \pi l^2$. The kth-nearest neighbor distance properties using chord length "distance" on this curved surface then appear analogous to those on a flat surface. There is, however, one important distinction. The relevant threshold w_0 for edge effects is in this case $w_0 = \pi(2R)^2$, where R is the radius of the sphere. Since $\pi(2R)^2$ is exactly equal to the total surface area of the sphere, it is set to 1. Equation (IV.4) thus requires no corrections at all, and so the universality in (IV.7) is exact.

— 4 — Curved Surfaces

Let us turn to the case of a surface with true curvature, where the distance is now defined in terms of a metric, i.e., along geodesics of the surface. Let us consider again a spherical surface. With l now representing arc length, the area of the spherical cap spanned by an arc originating at a pole of the sphere is given by

$$\begin{aligned} A(l)_{\text{sphere}} &= 2\pi R^2 \left[1 - \cos\frac{l}{R}\right] \\ &= 4\pi R^2 \sin^2\frac{l}{2R} \end{aligned}$$

If the total surface area, $4\pi R^2$, is normalized to 1,

$$\begin{aligned} A(l)_{\text{sphere}} &= \sin^2\sqrt{\pi} l, \qquad \text{so} \\ A^{-1}(w)_{\text{sphere}} &= \frac{1}{\sqrt{\pi}} \sin^{-1}\sqrt{w} \\ &\simeq \sqrt{\frac{w}{\pi}} \sum_{j=0}^{\infty} \frac{w^j}{2j+1} \frac{(2j)!}{2^{2j}(j!)^2}. \end{aligned} \tag{IV.8}$$

As in the case of the chord length "distance", this $A^{-1}(w)$ expression is exact everywhere for $0 \le w \le 1$. Equations (IV.3) and (IV.6) then require no corrections, and we find

$$\begin{aligned}\langle \tilde{D}_k(N) \rangle_{\text{sphere}} &= \sum_{j=0}^{\infty} \frac{1}{2j+1} \frac{(2j)!}{2^{2j}(j!)^2} \frac{(k+j-1/2)!}{(k-1/2)!} \frac{\sqrt{N}\, N!}{(N+j+1/2)!} \\ &= 1 + \frac{4k-7}{24N} + O\left(\frac{1}{N^2}\right). \end{aligned} \tag{IV.9}$$

Clearly, the k universality does not apply here.

Another sort of universality, however, is found when we turn to the more general case of an arbitrary closed surface, i.e., an abstract 2-D manifold with non-constant curvature and no boundary. For any surface, we may introduce a system of curvilinear coordinates u and v enabling us to write (at least piecewise) the differential length element ds in the conformal, orthogonal form (Eisenhart, 1909):

$$ds^2 = f(u,v)\,[du^2 + dv^2]. \tag{IV.10}$$

The Gaussian curvature $K(u,v)$ of the surface may then be expressed in terms of the function $f(u,v)$ by

$$K = \frac{1}{2f^3}\left[\left(\frac{\partial f}{\partial u}\right)^2 + \left(\frac{\partial f}{\partial v}\right)^2 - f\frac{\partial^2 f}{\partial u^2} - f\frac{\partial^2 f}{\partial v^2}\right]. \tag{IV.11}$$

What is $A(l)$ on this surface? In order to know this, we must first find out what are the manifold's geodesic lines. For ds given by (IV.10), we may use the geodesic equation:

$$\frac{d^2u}{ds^2} + \frac{1}{2f}\frac{\partial f}{\partial u}\left[\left(\frac{du}{ds}\right)^2 - \left(\frac{dv}{ds}\right)^2\right] + \frac{1}{f}\frac{\partial f}{\partial v}\frac{du}{ds}\frac{dv}{ds} = 0. \tag{IV.12}$$

Let us expand u and v as functions of distance s from an initial point, along a fixed geodesic:

$$\begin{aligned} u(s) &= u_0 + s u_0' + \frac{s^2}{2}u_0'' + \cdots \quad \text{and} \\ v(s) &= v_0 + s v_0' + \frac{s^2}{2}v_0'' + \cdots, \end{aligned} \tag{IV.13}$$

where $u_0 \equiv u(0)$, $u_0' \equiv u'(0)$, etc., and likewise for v. Then, expanding $f(u,v)$ in terms of u and v and substituting (IV.13),

$$\begin{aligned} f(u,v) &= f(u_0,v_0) + s\left[u_0' f_u(u_0,v_0) + v_0' f_u(u_0,v_0)\right] + s^2\left[\frac{u_0''}{2}f_u(u_0,v_0) + \frac{v_0''}{2}f_v(u_0,v_0)\right. \\ &\quad \left. + \frac{(u_0')^2}{2}f_{uu}(u_0,v_0) + u_0'v_0' f_{uv}(u_0,v_0) + \frac{(v_0')^2}{2}f_{vv}(u_0,v_0)\right] + O(s^3), \end{aligned}$$

where subscripts on f denote partial derivatives.

Using (IV.10) and (IV.12), we can solve for all but three of the coefficients in (IV.13). Let us choose u_0, v_0 and u_0' to be these three. Now, consider the area $A(l)$ about the point (u_0, v_0). For ds given in (IV.10), the differential surface element will be $d\mu = f\,du\,dv$, so:

$$\begin{aligned} A(l) &= \int f\,du\,dv \\ &= \int f\,J\left(\frac{u,v}{s,u_0'}\right) ds\,du_0' \\ &= \int f\,\left|\frac{\partial u}{\partial s}\frac{\partial v}{\partial u_0'} - \frac{\partial u}{\partial u_0'}\frac{\partial v}{\partial s}\right| ds\,du_0'. \end{aligned} \tag{IV.14}$$

The limits of integration over s are 0 and l; the limits of integration over u_0', which may be found from (IV.10), are $-\sqrt{1/f(u_0,v_0)}$ and $\sqrt{1/f(u_0,v_0)}$. In evaluating the Jacobian some care must be taken, as a sign ambiguity allows two solutions for the coefficients in (IV.13). $A(l)$ will be the sum of (IV.14) evaluated at each of the two solutions, ultimately causing all odd powers of l to vanish.

The result, after lengthy algebraic manipulations, may be written as

$$A(l) = \pi l^2\left[1 - \frac{l^2}{24}\,\frac{1}{f^3}\,(f_u^2 + f_v^2 - f f_{uu} - f f_{vv}) + O(l^4)\right], \tag{IV.15}$$

where in order to avoid cluttering the notation, we have omitted the (u_0, v_0) arguments at which all functions are to be evaluated. For the leading correction term in $A(l)$ given by (IV.15), we recognize the expression (IV.11) for the Gaussian curvature K. In retrospect, this is not surprising: by symmetry, the correction series to πl^2 can contain only even powers of l, and if we consider $A(l)$ as a geometric expansion about a flat space approximation, K will be the only scalar quantity with dimensions of l^{-2}.

With some perseverance, one may carry (IV.15) to higher orders in l, and the coefficients of the expansion turn out to be

$$\begin{aligned} A(l) &= \pi l^2\left[1 - \frac{l^2}{12}\,K + \frac{l^4}{720}\big(2K^2 - 3\nabla^2 K\big)\right. \\ &\qquad \left. - \frac{l^6}{161280}\big(8K^3 - 3[10(\nabla K)^2 + 14K\nabla^2 K - 5\nabla^4 K]\big) + O(l^8)\right], \end{aligned} \tag{IV.16}$$

where ∇ is the gradient operator. We thus have a series expansion giving the area of a disc on any 2-D surface, in a form that depends only on intrinsic quantities, *i.e.*, not on the choice of coordinate system.

We may invert (IV.16) to obtain the power series

$$\begin{aligned} A^{-1}(w) &= \sqrt{\frac{w}{\pi}}\left[1 + \frac{K}{24\pi}\,w + \frac{9K^2 + 4\nabla^2 K}{1920\pi^2}\,w^2 + \frac{15K^3 + 14K\nabla^2 K - 2(\nabla K)^2 + \nabla^4 K}{21504\pi^3}\,w^3\right. \\ &\qquad \left. + O(w^4)\right]. \end{aligned} \tag{IV.17}$$

Take as an example the special case of a spherical surface, where the Gaussian curvature is a constant $K = 1/R^2$, or $K = 4\pi$ for a unit surface. All derivatives of K then vanish, leaving

$$A^{-1}(w)_{\text{sphere}} = \sqrt{\frac{w}{\pi}}\left[1 + \frac{w}{6} + \frac{3w^2}{40} + \frac{5w^3}{112} + O(w^4)\right],$$

from which we recover the first few terms of our earlier result (IV.8).

Given an expression for $A^{-1}(w)$ on a general 2-D surface, we may find $\langle \tilde{D}_k(N) \rangle$ using (IV.3) and (IV.6). Let us consider the average $\int \langle \tilde{D}_k(N) \rangle \, d\mu$ over the entire surface, which is found directly from the average of the series coefficients in (IV.17). Now examine the $O(w)$ term in the series. By the Gauss-Bonnet theorem (EISENHART, 1909), $\int K \, d\mu = 2\pi\chi$ on any closed surface, where χ is the Euler characteristic of the surface, a topological invariant. We therefore obtain, up to leading corrections,

$$\begin{aligned} \int A^{-1}(w) \, d\mu &= \sqrt{\frac{w}{\pi}} \left[1 + \frac{\chi}{12} w + O(w^2) \right], \qquad \text{giving} \\ \int \langle \tilde{D}_k(N) \rangle \, d\mu &= \frac{\sqrt{N} \, N!}{(N+1/2)!} \left[1 + \frac{\chi}{12} \, \frac{k+1/2}{N+3/2} + O\left(\frac{1}{N^2} \right) \right] \\ &= 1 + \frac{\chi(2k+1) - 9}{24N} + O\left(\frac{1}{N^2} \right). \end{aligned} \tag{IV.18}$$

We thus discover another, very different sort of universality from the one we had in the case of flat space. To $O(1/N)$, the finite size scaling law for kth-nearest neighbor distances depends only on the surface's topology, and not on its detailed properties.

The Euler characteristic χ for a surface is related to its genus g by $\chi = 2(1-g)$. Taking the torus as one example, $g = 1$, so $\chi = 0$ and the k-dependence in (IV.18) once again disappears, at least to $O(1/N)$. This is to be expected: a flat space with periodic boundary conditions has, after all, the topology of a torus. And conversely, because of the topological invariance, all tori behave like flat space to $O(1/N)$. Taking the spherical surface as another example, $g = 0$, so $\chi = 2$ and we recover from (IV.18) the result (IV.9).

The properties of $\int \langle \tilde{D}_k(N) \rangle \, d\mu$ are far less clear at higher order in $1/N$. Using the divergence theorem and integration by parts, we may obtain from (IV.17)

$$\int A^{-1}(w) \, d\mu = \int \sqrt{\frac{w}{\pi}} \left[1 + \frac{K}{24\pi} \, w + \frac{3K^2}{640\pi^2} \, w^2 + \frac{15K^3 + 16K\nabla^2 K}{21504\pi^3} \, w^3 + O(w^4) \right] d\mu. \tag{IV.19}$$

If we looked only at terms up through $O(w^2)$, we might believe that this series is simply, by analogy with (IV.8), an expansion of $\int (2/\sqrt{K}) \sin^{-1} \sqrt{Kw/4\pi} \, d\mu$. Unfortunately, starting at $O(w^3)$ we see this is not true, since the contributions of curvature and its gradients do not all vanish in the average over the surface! Furthermore, even for terms in (IV.19) of the form $\int K^n \, d\mu$, at $n > 1$ there is no clear equivalent to the Gauss-Bonnet theorem; the theorem is a direct consequence of the integrand's linearity. Thus for a general 2-D surface, there does not appear to be a simplified form for the terms in $\int \langle \tilde{D}_k(N) \rangle \, d\mu$ beyond $O(1/N)$. More particularly, the *only* case in which $\int \langle \tilde{D}_k(N) \rangle \, d\mu$ would be independent of k beyond $O(1/N)$ is if the curvature is identically equal to 0, i.e., a flat surface.

Let us briefly consider the case of curved higher-dimensional manifolds. The calculation is now far more complicated, as it is no longer possible to write the metric tensor in a conformal form as we did in (IV.10). In addition, whereas in 2-D the only intrinsic scalar quantity describing curvature is the Gaussian curvature K, for $d > 2$ there are $d(d-1)(d-2)(d+3)/12$ different such quantities (WEINBERG, 1972). However, all of them except K itself have at least dimensions of

l^{-4}. It thus seems reasonable to conjecture that, as we argued in 2-D, the $O(l^2)$ correction term in $A(l)$ can only involve K. (Indeed, we have verified that this is true in 3-D.) In that case, we may rely on the example of the spherical surface — easily generalized to d dimensions — to provide us with the initial terms for a general manifold:

$$\begin{aligned} A(l) &= \frac{\pi^{d/2}}{(d/2)!}\, l^d \left[1 - \frac{d(d-1)}{d+2}\,\frac{K}{6}\, l^2 + O(l^4)\right], \qquad \text{or} \\ A^{-1}(w) &= \left[\frac{(d/2)!}{\pi^{d/2}}\right]^{1/d} w^{1/d} \left[1 + \frac{d-1}{d+2}\,\frac{K}{6}\left(\frac{(d/2)!}{\pi^{d/2}}\right)^{2/d} w^{2/d} + O(w^{4/d})\right]. \end{aligned} \tag{IV.20}$$

Note that $A^{-1}(w)$ now contains a series in $w^{2/d}$ rather than in w. Appropriately modifying (IV.3), it may then be shown that $\langle \tilde{D}_k(N) \rangle$ is in general given by a series in $1/N^{1/d}$ for odd d, and $1/N^{2/d}$ for even d.

Consider, finally, the average $\int A^{-1}(w)\, d\mu$ over the manifold. The higher dimensional generalization of the Gauss-Bonnet theorem (Nakahara, 1990) involves an integrand of $O(1/l^d)$, or $O(1/w)$. The leading correction term from (IV.20), $\int K\, d\mu$, therefore cannot be simplified further for $d > 2$; the only term that could possibly give rise to a topological invariant is the coefficient at $O(l^d)$, or $O(w)$. If d is odd, it is rather certain that no topological invariant will exist in the series. If d is even, the $O(w)$ term will first contribute to the $\int \langle \tilde{D}_k(N) \rangle\, d\mu$ series at $O(1/N)$ — as in 2-D, although at higher dimensions this will no longer be the leading correction term. While we cannot rule out the possibility of indeed obtaining a topological invariant at $O(1/N)$, the $O(w)$ term in $A^{-1}(w)$ is in general a complicated one involving many different curvature scalars, so this is far from obvious. It remains an open question.

— 5 — Regge Calculus

We have remarked that from a physical point of view it is natural, in the 2-D case, for the leading corrections in $A(l)$ to contain only the Gaussian curvature, as this represents the leading deviation from planarity. Consequently, only the mean curvature — or, using the Gauss-Bonnet theorem, the Euler characteristic χ — matters in the $O(1/N)$ term of $\int \langle \tilde{D}_k(N) \rangle\, d\mu$. We have seen using differential methods (geodesics) that this physical picture is indeed correct. These methods apply to a smooth surface. For polyhedral surfaces, which are not smooth, we may in fact obtain a similar result more easily, using the *non-differential* method of Regge calculus. Consider a polyhedron with a number of vertices, edges, and faces. Following the work of Regge (Regge, 1961) and others since then (Cheeger, Muller & Schrader, 1984), we observe that the curvature is concentrated at the vertices and is measured by a deficit angle: if θ_i is the sum of the angles incident on vertex i, the deficit angle at that vertex is $\Delta_i = 2\pi - \theta_i$. It may then be shown that the Gauss-Bonnet theorem, on polyhedra, reduces to Euler's relation $2\pi\chi = \sum_i \Delta_i$.

Let P be a polyhedron with a fixed number of vertices, and consider the problem of finding the finite size scaling $\int \langle \tilde{D}_k(N) \rangle\, d\mu$ on P. As $N \to \infty$, corrections to the flat space value about a given point $\mathbf{x}$ arise only when $\mathbf{x}$ is near one of the vertices, because only then can curvature

(i.e., the deficit angle) enter into the local calculation of $A(l)$ about $\mathbf{x}$. It is then sufficient to understand the corrections associated with one vertex at a time. Consider a vertex i. $A(l)$ receives a contribution from i that depends exclusively on the deficit angle Δ_i. For small deficit angles (small corrections to $A(l)$), a linear approximation may be used and this contribution will simply be proportional to Δ_i. Correspondingly, the leading corrections to $A(l)$ — and thus the $O(1/N)$ term in the finite size scaling series — will be proportional to Δ_i. Summing over all the vertices $i = 1, \ldots, N$, we find that the leading correction term in $\int \langle \tilde{D}_k(N) \rangle \, d\mu$ is indeed a linear function of χ, and we recover the topological invariant derived in the case of a smooth manifold.

A word of caution is necessary, however. It is tempting at this point to take the limit where P becomes a smooth manifold, thus recovering (IV.18). Unfortunately this will not work; a direct computation shows that the limit does not commute with the limit $N \to \infty$ taken above, and the coefficient thus obtained for the $O(1/N)$ term will not be the correct one.

— 6 — Conclusions

We have considered the finite size scaling of mean distances to neighboring sites, when N sites are distributed randomly and uniformly on a surface with no boundaries. When the surface is flat, we have found that the entire $1/N$ series describing the mean kth-nearest site distance is independent of k. This universality applies equally well to higher moments of the distances, and to Euclidean manifolds in dimensions greater than 2. For surfaces with curvature, while this general property is no longer valid, we have found that the leading correction term in the series, averaged over the surface, is a topological invariant. The scaling series thus depends, to $O(1/N)$, on the genus of the manifold but not on its other properties.

Although we have considered these universalities only for the moments of point-to-point distances, similar properties hold for higher order simplices such as areas of triangles associated with nearby points. The problem is thus a natural one to consider further in the context of random triangulations, foams or other physical problems (Itzykson & Drouffe, 1989) tightly connected to geometry.

Acknowledgments

We are grateful to J. Houdayer and E. Bogomolny for sharing with us their numerous and valuable insights on this topic, and to O. Bohigas for having introduced us to the problem. OCM acknowledges support from the Institut Universitaire de France.

Bibliography

Cheeger J., Muller W. & Schrader R. (1984): "On the Curvature of Piecewise Flat Spaces", *Comm. Math. Phys.*, **92**(3), pp. 405–454.

Eisenhart L.P. (1909): *A treatise on the differential geometry of curves and surfaces*. Boston: Ginn and Company.

Itzykson C. & Drouffe J. (1989): *Statistical Field Theory*. Cambridge: Cambridge University Press.

Nakahara M. (1990): *Geometry, Topology and Physics*. Bristol: A. Holger.

Regge T. (1961): "General Relativity without Coordinates", *Nuovo Cimento*, **19**, pp. 558–571.

Weinberg S. (1972): *Gravitation and Cosmology: principles and applications of the general theory of relativity*. New York: Wiley.

– Appendix A –

Computational complexity: P vs. NP

The inherent challenge of the TSP lies in the fact that no known algorithm can find the optimal tour of an arbitrary N-city instance in a number of steps polynomial in N. In computational complexity theory, this is what classifies it as a "hard" problem. An "easy" problem would be one that could be solved by a polynomial algorithm. The word "easy", of course, does not imply "fast": an $O(N^3)$ algorithm, at $N = 10{,}000$, could already involve immense running times! There is nevertheless a large conceptual difference between a problem that can be solved in no worse than polynomial time, and a problem that requires, say, exponential time.

Computer scientists formalize these concepts with the notion of P and NP. A problem belongs to class P if there exists an algorithm that solves the problem in a time growing polynomially, or slower, with the size N of the problem. A problem belongs to class NP if it is merely possible to test, in polynomial time, whether a certain "guessed" solution is indeed correct. Note that NP stands for "non-deterministic polynomial", and *not*, as is sometimes mistakenly thought, for "non-polynomial". P is in fact a subset of NP, since for any problem whose solution can be found in polynomial time, one can surely verify the validity of a potential solution in polynomial time.

Formally speaking, the classes P and NP apply only to *decision* problems, where the solution is simply a "yes" or a "no". It is however, generally possible to relate combinatorial optimization problems to associated decision problems. For the case of the TSP, instead of asking for the optimal tour length, we may phrase the associated TSP-*decision* as follows: is there a tour whose length is less than a given value B? It is relatively straightforward to see that the optimal TSP tour length can be found to arbitrary precision by executing a sequence of TSP-decisions (with bound B, say, decreasing in discrete steps), where the length of the sequence is polynomial in N (Johnson & Papadimitriou, 1985). The TSP-decision is thus said to be *polynomially reducible* to the TSP; this implies that the two problems, though not necessarily themselves equivalent, are at least of equivalent computational complexity.

In order to see that the TSP-decision is in NP, let us use the following more careful definition of the NP class (Cook, 1971). A decision problem belongs to NP if any instance calling for a "yes" response contains a *certificate*, itself of size polynomial in N, allowing the "yes" answer to be verified in polynomial time. For the TSP-decision, this certificate is simply a tour of length less than B: the certificate is of size N, and in $O(N)$ steps one may verify that it is a legal tour and satisfies the length bound. The certificate is therefore the "guess" that is tested. Of course, we cannot hope to verify a "no" answer in the same way as a "yes" answer; the only property characterizing a "no" instance is that *no certificate exists*. The fact that a decision problem is in NP merely means that we can confirm a "yes" instance if we happen to be given the right certificate.

In the case of the TSP-decision, one can go further. An important subset of the class of NP problems is the class of NP-complete problems. These are problems that are *as least as complex* as any other NP problem — in other words, an algorithm capable of solving an NP-complete problem could be mapped onto an algorithm capable of solving *any* NP problem, via a

polynomial encoding. (This is not quite the same thing as the notion of polynomial reducibility, as the encoding associates a single instance of one problem with a single instance of another, rather than with a sequence of instances.) It has been proven that the TSP-decision is NP-complete (PAPADIMITRIOU, 1977). Let us consider the implications of this. We have already noted that P $\subseteq$ NP, since a problem solvable in polynomial time has by definition a certificate (its solution!) verifiable in polynomial time. If it turns out that a polynomial algorithm could be found for solving the TSP-decision — or any other NP-complete problem — one would exist for all NP problems and we would have P = NP. No such polynomial algorithm has ever been found, and it is conjectured that none exists, so that P $\neq$ NP. Proving this conjecture, however, has been an open question in complexity theory since the 1970s.

Note that although it has become standard practice to refer to the TSP as being NP-complete, the more correct term is actually NP-*hard*. This is because, strictly speaking, the complexity classes P, NP and NP-complete are reserved for decision problems. NP-hard is the more general designation for problems to which NP-complete decision problems are polynomially reducible. The TSP-decision is NP-complete; the TSP itself, being of equivalent complexity, is thus NP-hard (see GAREY & JOHNSON, 1979; JOHNSON & PAPADIMITRIOU, 1985).

A different issue from that of *solving* an NP-hard problem in polynomial time is that of *approximating* its solution in polynomial time. For the Euclidean TSP, KARP (1977) developed the *fixed dissection algorithm*, which generates a tour using the "divide-and-conquer" strategy of partitioning space into subparts, solving the TSP for the cities within each subpart, and joining up these subtours in such a way as to form one large (non-optimal) tour through all cities. Karp proved that this algorithm can with high probability (prob. $\to 1$ as $N \to \infty$) give tours whose length is within $(1+\epsilon)$ times optimality, for arbitrary ϵ. The expected execution time of this algorithm is $O(N^2 \log N)$.

Karp's construction has three limitations: first, the heuristic does not *guarantee* the $1+\epsilon$ bound at finite N; second, the running time could in a worst-case situation be exponential; third, the method only applies to instances from the random ensemble, with a uniform distribution of cities. Recent work by ARORA (1997) appears to have resolved these three difficulties in the case of the Euclidean TSP. Using a different partitioning scheme, Arora's algorithm finds a guaranteed $(1+\epsilon)$-approximation in $O\big((\log N)^{O(\sqrt{d}/\epsilon)^{d-1}} N\big)$ time, for arbitrary instances in d dimensions. This sort of method, unfortunately, cannot easily be generalized to other problems. It has in fact been proven (ARORA, LUND, MOTWANI, SUDAN & SZEGEDY, 1992) that if P $\neq$ NP, $(1+\epsilon)$-approximation algorithms cannot exist for all NP-hard problems. The fact that one has been found for the Euclidean TSP is quite remarkable, and suggests that further refinement of the NP-hard classification may be necessary.

– Appendix B –

Outline of a self-averaging proof

The original proof of self-averaging in the Euclidean optimal tour length $L_E(N,d)$, by Beardwood, Halton & Hammersley (1959), is quite technical. There is however a more accessible proof, by Karp & Steele (1985), using the *fixed dissection algorithm* (Karp, 1977) mentioned in Appendix A. This algorithm generates a (non-optimal) tour by dividing the space into subparts, finding the optimal subtour within each subpart, and connecting up all of these subtours (minus one link, in each case) to form one large tour. Karp and Steele also took advantage of a lemma (proved by induction), stating that in the d-dimensional unit hypercube there must always exist a tour of length less than $dN^{1-1/d} + \delta_d N^{1-1/(d-1)}$ (δ_d is a constant depending only on d).

They then considered a Poisson process placing points with unit intensity in the hypercube $[0,t]^d$, and looked at the expectation value $F(t)$ of the optimal tour length through these points. If $\langle L_E(N,d)\rangle$ is the expected optimal tour length through N points in the unit hypercube, t times this quantity will be the expected optimal tour length through N points in $[0,t]^d$. From the Poisson distribution, we thus obtain:

$$F(t) = \sum_{N=0}^{\infty} e^{-t^d}\, \frac{t^{dN}}{N!}\, t\, \langle L_E(N,d)\rangle. \tag{B.1}$$

Now, the fixed dissection algorithm bounds the optimal tour length through the cities placed by the Poisson process. This upper bound is equal to the lengths of the various subtours plus the length of the large circuit through space needed to connect the subtours. Bounding the latter length using the lemma above, and partitioning space into m^d equal subparts, we then obtain the expectation value bound:

$$F(t) \;\le\; m^d F(t/m) + t\big(d(m^d)^{1-1/d} + \delta_d (m^d)^{1-1/(d-1)}\big), \text{ or} \tag{B.2}$$

$$\frac{F(t)}{t^d} \;\le\; \frac{F(t/m)}{(t/m)^d} + \frac{d}{(t/m)^{d-1}} + \delta_d\, \frac{m^{-1/(d-1)}}{(t/m)^{d-1}}. \tag{B.3}$$

Clearly $F(t)$ is monotone increasing, and because of the bounding lemma, $F(t)/t^d$ must be bounded. As $t \to \infty$, then, for fixed m, (B.3) implies that $F(t)/t^d$ approaches a limiting value, which we shall call $\beta_E(d)$. From (B.1), substituting $u = t^d$,

$$\lim_{u\to\infty} \sum_{N=0}^{\infty} e^{-u}\, \frac{u^N}{N!}\, \frac{\langle L_E(N,d)\rangle}{u^{1-1/d}} = \beta_E(d). \tag{B.4}$$

From the Tauberian theorem (see Karp & Steele, 1985, p. 187) for the Poisson distribution, finally, this gives:

$$\lim_{N\to\infty} \frac{\langle L_E(N,d)\rangle}{N^{1-1/d}} = \beta_E(d). \tag{B.5}$$

The statement is in fact stronger than it appears: Karp and Steele are also able to bound the variance of $L_E(N,d)$ using an inequality due to Efron & Stein (1981), giving $\mathrm{Var}[L_E(N,d)/N^{1-1/d}] = O(N^{-1})$. We now know, therefore, not only that the distribution of $L_E(N,d)/N^{1-1/d}$ becomes increasingly sharply peaked at large N, but also that its width goes to zero as $1/\sqrt{N}$, *i.e.*, as a Gaussian.

– Appendix C –

Estimates for a non-uniform distribution of cities in space

We have seen that the optimal tour length is self-averaging, in the ensemble of cities independently and uniformly distributed in Euclidean space. With probability 1, the random variable $L_E(N,d)/N^{1-1/d}$ tends to a constant $\beta_E(d)$ in the $N \to \infty$ limit. The original proof of this result, due to Beardwood, Halton & Hammersley (1959), is however more general and deals with an arbitrary density $\rho(\mathbf{x})\, d\mathbf{x}$ of independently distributed cities in space. The general statement of self-averaging is that, with probability 1,

$$\lim_{N\to\infty} \frac{L_E(N,d)}{N^{1-1/d}} = \beta_E(d) \int [\rho(\mathbf{x})]^{1-1/d}\, d\mathbf{x}, \tag{C.1}$$

where $\beta_E(d)$ is independent of the density, and thus equal to its value for a uniform distribution. What this means in practice is that the numerical value of $\beta_E(d)$ is of relevance for cities distributed independently in space with *any* distribution.

In this appendix we propose to use this result to give an *a priori* estimate of L_E for any instance. Note, first, that the factor $\int [\rho(\mathbf{x})]^{1-1/d}\, d\mathbf{x}$ applies just as well to quantities such as the distance $D_1(N,d)$ between nearest neighbors. For an arbitrary Euclidean instance at large N, then, we may estimate the optimal tour length $L_E^{\text{non-uniform}}$ for that instance as:

$$L_E^{\text{non-uniform}} \approx D_1^{\text{non-uniform}} \frac{\langle L_E^{\text{uniform}} \rangle}{\langle D_1^{\text{uniform}} \rangle} \tag{C.2}$$

where $D_1^{\text{non-uniform}}$ is the mean nearest neighbor distance for the instance we are considering. $\langle L_E^{\text{uniform}} \rangle$ is given by its value in the Euclidean uniform ensemble, and $\langle D_1^{\text{uniform}} \rangle$ is the ensemble average of nearest neighbor distances. Both of these quantities, as functions of N and d, have been calculated in Chapter II. Note that there is no need to use units where volume equals 1 in the non-uniform instance under consideration; our use of $D_1^{\text{non-uniform}}$ in (C.2) will take into account any volume scaling.

Let us try this out on a TSP instance well known in the operations research literature, the AT&T-532 instance.[1] This consists of 532 sites belonging to the U.S. telecommunications company. For clear commercial reasons, the distribution of sites in the instance follows roughly the population density of the country, and thus is highly non-uniform. The instance was first solved to optimality (see Figure C–1) by Padberg & Rinaldi (1987), and in their units, the optimum was found to be 27,868. How closely does our approach predict this result? In their units, once again, the mean distance between a city and its nearest neighbor can be measured as 36.205. Using the expressions given in Chapter II, we would then estimate:

[1] AT&T-532 is one of the many instances found in TSPLIB, a library assembled by Reinelt (1991) and available from *http://softlib.rice.edu/softlib/tsplib*.

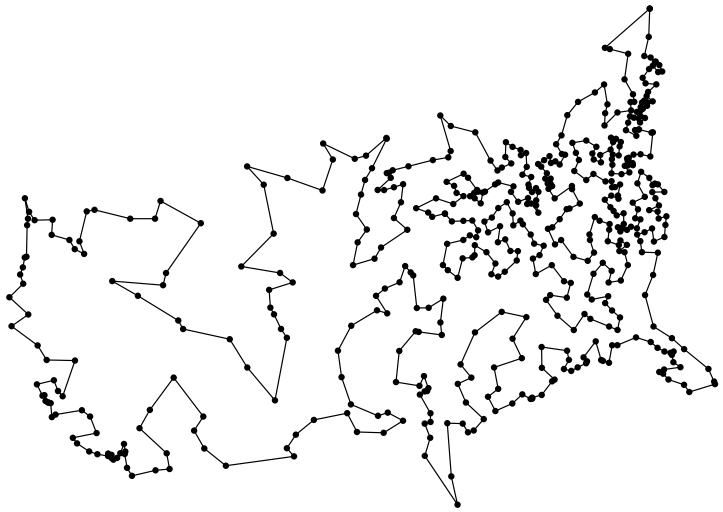

Figure C–1: Optimal tour in the (non-uniform) AT&T-532 instance.

$$L_E^{\text{non-uniform}} \approx D_1^{\text{non-uniform}} \frac{\langle L_E^{\text{uniform}} \rangle}{\langle D_1^{\text{uniform}} \rangle} \tag{C.3}$$

$$\approx D_1^{\text{non-uniform}} \times \beta_E(2) \times \left(1 - \frac{0.0171}{N} - \frac{1.048}{N^2}\right) \times 2N \tag{C.4}$$

$$\approx 27{,}427 \tag{C.5}$$

which is within 1% of the actual optimum.

It may seem surprising that our finite size scaling law *given periodic boundary conditions* provides such a good estimate for a tour length calculated given open boundary conditions. There are two reasons for this agreement. First of all, boundary effects are involved in $D_1^{\text{non-uniform}}$. (C.2) incorporates these effects into the $L_E^{\text{non-uniform}}$ estimate, and although they may not be precisely the correct effects for the optimum tour length, they are undoubtedly close. Second of all, $\beta_E(2)$ itself does not depend on the boundary conditions chosen. For a discussion of why this is so, the reader is referred to JAILLET (1993), who has proven that open and periodic boundary conditions give the same $\beta_E(d)$ value.

– Appendix D –

Numerical methodology

In this appendix we discuss the algorithms used to obtain our simulation results. For both the Euclidean and random link TSPs, we have performed runs at instance sizes $N = 12$, 13, 14, 15, 16, 17 using the LIN & KERNIGHAN (1973) heuristic (LK), and at $N = 30$ and $N = 100$ using the *Chained Local Optimization* heuristic (CLO) of MARTIN, OTTO & FELTEN (1991).

LK heuristic

The LK algorithm may be sketched as follows. The kernel of the algorithm is what might be termed LK-*search*. We start off with an initial (non-optimal) tour. LK-search takes an arbitrary starting city — for convenience, we generally start with the top city on a fixed "list" in our program. Call this city i_0. Now pick a starting direction — either "forwards" or "backwards" — and call i_1 the next city in that direction along the initial tour. Call l_1 the link between i_0 and i_1, and remove that link. We will now attempt to reconnect i_1 to a new city i'_1 that is somewhere else on the tour, resulting in the situation shown schematically in Figure D–1. Let l'_1 be the link to that new city i'_1. How do we choose i'_1, and thus l'_1? i'_1 is the nearest neighbor to i_1 that was not already connected to i_1 in the initial tour. There is another very important requirement for i'_1: it must be such that $l'_1 < l_1$. This it the *gain criterion*, and is applied at all stages of LK-search. If the gain criterion cannot be met, the search using this i_0 and l_1 is abandoned. If the gain criterion is met, however, the search deepens.

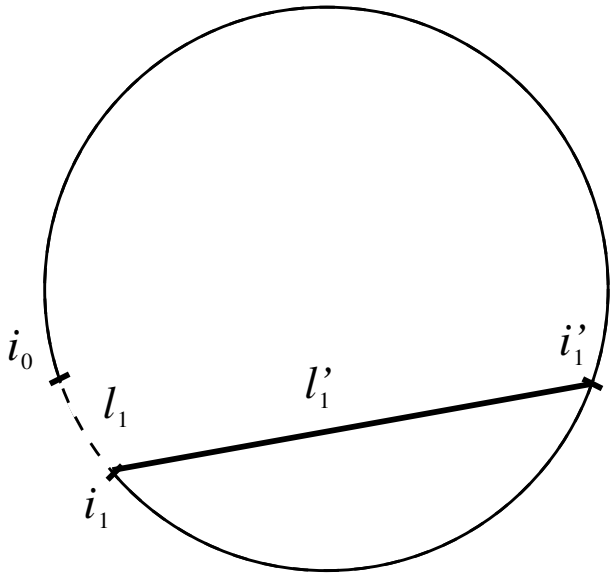

Figure D–1: One step of LK-search, showing the removed link l_1 (dashed line) and added link l'_1 (bold line).

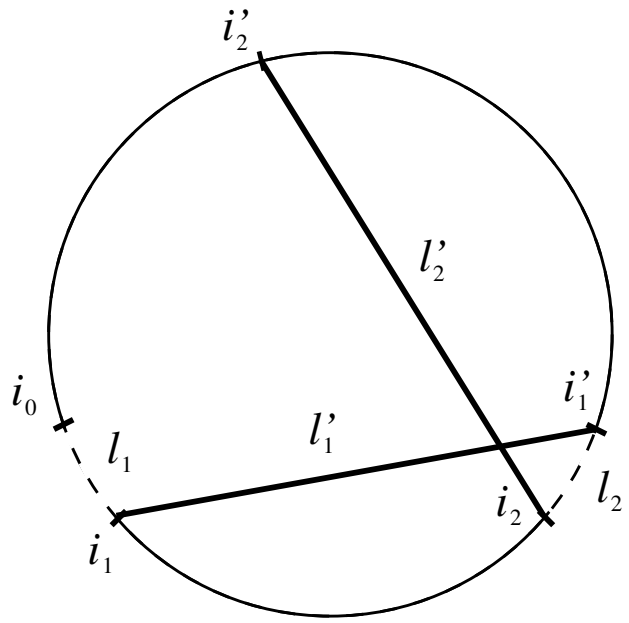

Figure D–2: LK-search at step 2, showing tadpole with two links removed and two links added.

At this point we have what might be called a "tadpole graph" (the more technical term is a *one-tree*). We now pick one of the two cities next to i_1' along the initial tour, call it i_2, and remove the link l_2 joining them. Which of the two possibilities do we use for i_2, and thus l_2? We pick i_2 such that, if i_2 were then to be connected to the dangling end (i_0), a single closed tour would result. Now continue as before. We attempt to reconnect i_2 to a new city i_2', resulting in the situation in Figure D–2. l_2' is the link to the new city, and that city, as before, is the nearest neighbor that was not an adjoining member of the original tour. The gain criteria get stronger as the search gets deeper: the total gain must remain positive, so we must have $l_1' + l_2' < l_1 + l_2$, or else the search aborts. LK-search continues recursively in this way, with the vertex of the "tadpole" hopping around with the end point i_0 staying fixed, until the gain criteria force it to abort. There is, however, an important additional part to the gain criteria. At each step m, whilst i_m is being chosen in such a way as to make it possible to close up the tour, the length of that closed tour is recorded. If at any step of LK-search the tadpole length for that step is longer than the best closed tour recorded so far, the search terminates.

When a search terminates, one of two things can happen. If the best recorded tour is shorter than the initial tour, the entire process begins again but using that tour as the new one.[1] If, on the other hand, the best recorded tour is not shorter than the initial tour, we recommence with the alternate choice of i_1: if the starting direction was "forwards", now we try "backwards", and vice versa. When, however, both choices of i_1 have been exhausted, we select a new i_0 by advancing to the next city on the list. The algorithm, in our implementation, terminates when

[1] This will happen, notably, if the search is abandoned because of the second part of the gain criteria (closed-up tour shorter than tadpole), as the first part of the gain criteria ($\sum_i l_i' < \sum_i l_i$) requires that a tadpole always be shorter than the initial tour.

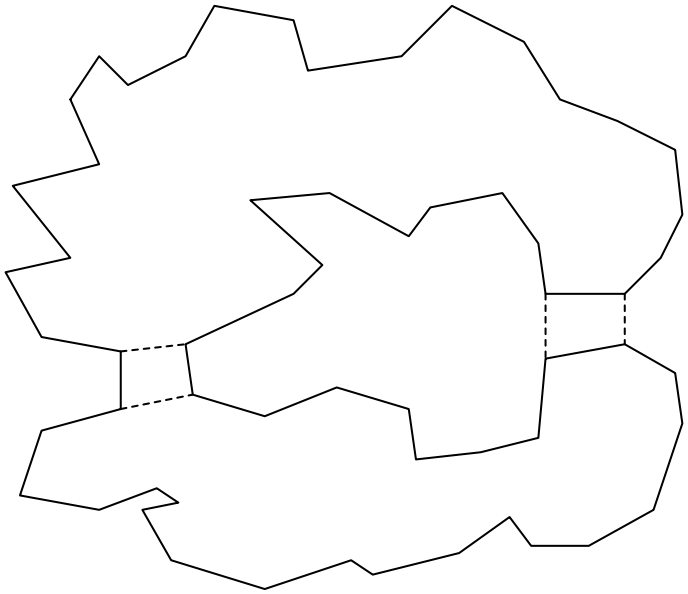

Figure D–3: *Double-bridge* change executed by CLO heuristic, generated by removing links shown with dashed lines, and reconnecting them differently.

an entire "pass" of N starting cities i_0 results in no improvement.

CLO heuristic

CLO is a stochastic algorithm, combining large Monte Carlo jumps with embedded LK local search. In our implementation, it works as follows. A (non-optimal) initial tour is optimized using LK. This brings it to an LK local minimum. An attempt is then made to modify the tour by a random 4-change (4 bonds disconnected and then reconnected differently), which would not be accessible by the kind of sequential changes performed by LK-search. The 4-change in question is known as a *double-bridge* kick, shown in Figure D–3: one 2-change disconnects the tour, and the other 2-change reconnects the two parts in a different place. After this 4-change is carried out, LK is again used to optimize the resulting tour. If the new local optimum is better than the previous one, the attempt succeeds, and CLO iterates the random 4-change procedure from the new tour. If not, the attempts fails and CLO tries another random 4-change on the old tour. This continues for a fixed number of steps. The idea is that, by directing its search through state space, CLO performs better than simply running LK from random starts an equivalent number of times.

Use of heuristics

For instance sizes in the range $12 \leq N \leq 17$, we used the LK heuristic. Our method consisted of two parts: first, we used a testbed of instances to estimate the systematic bias in LK arising from the fact that LK does not always find the true optimum; second, we did our "production runs" to determine the actual optimal tour lengths.

In order to determine the systematic bias, we performed the following procedure on values of N from 12 through 17: (i) we generated a testbed of 200 random instances; (ii) for each random instance, we generated 100 different random starting tours; (iii) for each random starting tour we ran the LK heuristic. Based on the assumption that the best LK-opt obtained over 100 random starting tours was indeed the true optimum for that instance, we calculated the expected bias per instance. We then averaged this bias over the testbed of instances, obtaining an estimate for each value of N.

For our production runs, we then performed the following procedure on values of N from 12 through 17: (i) we generated 100,000 random instances; (ii) for each random instance, we generated 10 different random starting tours; (iii) for each random starting tour we ran the LK heuristic. We then took, for each instance, the best LK-opt obtained over the 10 random starting tours, and averaged over all instances, obtaining an estimate of the mean optimum length for each value of N.

For instance sizes $N = 30$ and $N = 100$, we used the CLO heuristic with 10 steps. The method and procedures were identical to those for LK, apart from the numbers for the production runs: for $N = 30$ we generated 8,000 random instances and 5 random starting tours per instance; for $N = 100$ we generated 1,200 random instances and 20 random starting tours per instance.

Other operational issues

The choice of number of random starting tours per instance in our production runs was motivated by the need to keep the systematic bias to a minimum. For the LK runs, using 10 random starting tours per instance allowed us to have in the worst case, at $N = 17$, a bias of 1 part in 200,000 for the Euclidean TSP and 1 part in 170,000 for the random link TSP. This bias was therefore negligible at the level of numerical precision used (4 decimal places). For the CLO runs at $N = 30$, using 5 random starting tours per instance was sufficient to give us an estimated bias of under 1 part in 400,000 for the Euclidean TSP, and under 1 part in 800,000 for the random link TSP. We suspect these figures — particularly the latter — might be too good to be true, so it is possible that our estimate is corrupted somewhat by instances in the testbed where the true optimum was never found. (Even if our estimate of the bias is off by an order of magnitude, though, it will still be negligible compared with our statistical error bar, which at $N = 30$ is about 1 in 3,000 for both Euclidean and random link.) For the CLO runs at $N = 100$, using 20 random starting tours per instance gave an estimated bias of 1 part in 120,000 for Euclidean and 1 part in 60,000 for random link — again, comfortably below the statistical error of 1 in 6,000 for Euclidean and 1 in 2,000 for random link.

Finally, let us note a few further details concerning our code. Timings for our $d = 2$ random link simulations are given in Table D–1; as can be seen we ran on four different machines, two of them Dec Alphas and two of them Sun SPARCs. The code itself is written in C, and is a heavily modified version of the CLO package developed by Steve Otto (otto@cse.ogi.edu) and Robert Prouty (prouty@cse.ogi.edu). In the code, cities are placed on a 10,000 × 10,000 grid, and distances are rounded to the nearest integer. We confirmed that the mesh was fine enough for roundoff error to be negligible to within the 4-decimal-place precision of our numerical data. The random number generator used, from the CLO package, is a standard linear congruential

Table D–1: Simulation timings for $d = 2$ random link. Instance sizes $N = 12$ through $N = 17$ used LK, $N = 30$ and $N = 100$ used 10-step CLO. # runs indicates number of instances times number of random starts per instance.

N	LK/CLO	# runs	Machine	CPU time (secs)
12	LK	500,000	Dec Alpha200	7133.4
12	LK	500,000	Sun Ultra1	22001.2
13	LK	500,000	Dec Alpha200	10772.0
13	LK	500,000	Sun Ultra1	26025.2
14	LK	500,000	Sun Sparc100	32755.3
14	LK	500,000	Sun Sparc100	32420.9
15	LK	500,000	Dec Alpha200	10724.9
15	LK	500,000	Dec AS1000	9357.5
16	LK	500,000	Sun Sparc100	41459.0
16	LK	500,000	Sun Sparc100	38795.3
17	LK	500,000	Sun Ultra1	43508.3
17	LK	500,000	Sun Sparc100	46584.2
30	CLO	5,000	Sun Ultra1	10801.6
30	CLO	5,000	Sun Sparc100	11717.5
30	CLO	5,000	Sun Sparc100	11621.3
30	CLO	5,000	Sun Ultra1	11163.9
30	CLO	5,000	Sun Sparc100	10773.3
30	CLO	5,000	Sun Ultra1	10569.0
30	CLO	5,000	Sun Sparc100	11558.9
30	CLO	5,000	Sun Sparc100	11287.7
100	CLO	6,000	Dec Alpha200	19694.6
100	CLO	6,000	Dec AS1000	17982.2
100	CLO	6,000	Sun Sparc100	71671.1
100	CLO	6,000	Sun Sparc100	71137.9

algorithm working as follows:

```
#define MASK ( 0x7fffffff )
#define MULT 1103515245
#define ADD 12345
#define TWOTO31 2147483648.0
AAA = MULT & MASK;
BBB = ADD & MASK;
```

where & is the binary AND operator. The routine then updates the integer quantity randx to be (AAA*randx + BBB) & MASK, and returns a double-precision variable randx/TWOTO31.

– Appendix E –

Another numerical study of $\beta(d)$

Subsequent to our work, JOHNSON, MCGEOCH & ROTHBERG (1996) have used a variant of our approach to confirm the values $\beta_E(2) = 0.7120 \pm 0.0002$ and $\beta_E(3) = 0.6979 \pm 0.0002$ (to higher precision) and to extend the results to $d = 4$. Let us briefly summarize their work, and the differences between our approach and theirs.

Our finite-size scaling analysis, presented in Chapter II, uses data points at values of N from $N = 12$ to $N = 17$, as well as the points $N = 30$ and $N = 100$. Johnson *et al.*, making use of our scaling law (II.2), work with larger instances sizes — $N = 100$, $N = 316$ and $N = 1{,}000$ — where a better fit may presumably be obtained. This requires much more powerful computational resources than our own, but also good algorithms and efficient methods for reducing statistical error to a minimum (so as to avoid too large a number of time-consuming runs).

When one wishes to estimate $\langle L_E(N,d)\rangle$, the most direct way is to take the numerical average $\overline{L_E(N,d)}$ over a sample of K randomly chosen instances. This estimator has an expected statistical error $\sigma(K) = \sigma_{L_E}/\sqrt{K}$, where σ_{L_E} is the instance-to-instance standard deviation of L_E. One can improve on this, however. Let us instead measure $\overline{L_E(N,d) - L^*(N,d)} + \overline{L^*(N,d)}$. If L^* is a quantity closely correlated with L_E, $L_E - L^*$ will have a substantially lower variance than L_E itself; if in addition L^* has a small variance (or better yet, if we can measure the ensemble average $\langle L^*\rangle$ exactly) then we will ultimately obtain a measurement of $\langle L_E\rangle$ to higher accuracy than with the direct estimator $\overline{L_E}$. This is the method we have adopted in our own work. For L^*, we use $\lambda(L_1 + L_2)/2$, where λ is a parameter, L_1 is the mean distance between nearest neighboring points in an instance, and L_2 is the mean distance between second-nearest neighboring points in an instance. The great advantage of this approach is that an analytical expression can be found giving $\langle L_1 + L_2\rangle/2$ exactly, so the *only* statistical error is due to the estimator $\overline{L_E - \lambda(L_1 + L_2)/2}$. λ can be chosen so as to minimize the variance of this estimator: it is relatively simple to show that the value λ^* minimizing λ is given by

$$\lambda^* = \frac{\langle L_E(L_1 + L_2)/2\rangle - \langle L_E\rangle\langle (L_1 + L_2)/2\rangle}{(\sigma_{(L_1+L_2)/2})^2} \tag{E.1}$$

where $\sigma_{(L_1+L_2)/2}$ is simply the instance-to-instance standard deviation of $(L_1 + L_2)/2$. Notice that λ^* is in fact a measure of how closely correlated L_E and $(L_1 + L_2)/2$ are — if they are perfectly correlated, λ^* will be equal to 1, and if they are uncorrelated, λ^* will be equal to 0. In practice, we have found that the degree to which they are correlated varies little in N. For the $d = 2$ Euclidean case, we found $\lambda^* \approx 0.75$ at $N = 15$; keeping λ^* fixed to this value over all N, the variance reduction ranged from 0.38 at $N = 15$ to 0.43 at $N = 100$. We thus succeeded in reducing the overall statistical error $\sigma(K)$ by a factor of between 2.33 and 2.63. (For the random link TSP, the same method reduced the error by a factor of between 2.69 and 3.08.) The relative insensitivity of the variance reduction scheme to N suggests that correlations between L_E and $(L_1 + L_2)/2$ are stable at large N; a related property has been seen in the random link case in Chapter III (see Figure III–5).

In the work of JOHNSON, MCGEOCH & ROTHBERG (1996), a somewhat different procedure is used for reduction of variance. The emphasis of their article is on analyzing the *Held-Karp* lower bound, the solution to an integer programming relaxation of the TSP (HELD & KARP, 1970). Thus instead of measuring $\langle L_E \rangle$ using $\overline{L_E - L^*} + \overline{L^*}$, they use the estimator $\overline{L_E/L_{HK}}\ \overline{L_{HK}}$, where L_{HK} is the Held-Karp bound. The advantage of using this quantity is that it turns out to be correlated *very* closely with L_E, so that L_E/L_{HK} has extremely low instance-to-instance fluctuations. (To compare: at $N = 100$ in 2-D, our estimator $L_E - L^*$ leads to a statistical error for sample size K of about $1.3\%/\sqrt{K}$, whereas theirs leads to a statistical error of about $0.3\%/\sqrt{K}$.)[1] Furthermore, in measuring L_E/L_{HK} Johnson *et al.* used exact codes for the $N = 100$ and $N = 316$ instances, so there is no systematic error in these measurements (for $N = 1{,}000$ they used the same large N heuristic as we used for $N = 30$ and $N = 100$). For the second part of the estimator, $\overline{L_{HK}}$ itself, they used a rapid approximate method — though they were able to quantify the systematic bias inherent in this approximation (about 0.005%) on the basis of a testbed of instances where L_{HK} had already been found exactly. They then corrected for this systematic bias.

Note that while in our measurements, $\langle L^* \rangle$ is calculated exactly over the ensemble and the statistical error is due exclusively to the fluctuations of $L_E - L^*$, in the measurements of Johnson *et al.*, $\overline{L_{HK}}$ involves a significant statistical error whereas $\overline{L_E/L_{HK}}$ involves an almost negligible error. It is instructive to compare our $N = 100$ results with theirs (this is the only 2-D instance size we have in common with them). Our data, on the basis of 6000 instances, give $\langle L_E(100,2) - L^*(100,2) \rangle = 2.4342 \pm 0.0012$ and $\langle L^*(100,2) \rangle = 4.6934$, so $\langle L_E(100,2) \rangle = 7.1276 \pm 0.0012$. Johnson *et al.* obtain $\langle L_E(100,2)/L_{HK}(100,2) \rangle = 1.005542 \pm 0.000027$ on the basis of 13,957 instances, and $\langle L_{HK}(100,2) \rangle = 7.0897 \pm 0.0006$ on the basis of 98,246 instances, giving $\langle L_E(100,2) \rangle = 7.1290 \pm 0.0008$. We are unaware of the running times for the latter results, and thus cannot compare the relative efficiency of the two approaches. It is clear, however, that our resources could not permit instance sizes much in excess of $N = 100$, whereas their tests went up to $N = 1{,}000$.

There is, interestingly, a further bias in the estimator used by Johnson *et al.* that they have not noted. This arises from the fact that, of course, $\langle L_E \rangle \neq \langle L_E/L_{HK} \rangle \langle L_{HK} \rangle$. It is relatively straightforward to calculate this bias at large L_E and L_{HK}:

$$\left\langle \frac{L_E}{L_{HK}} \right\rangle \langle L_{HK} \rangle \approx \langle L_E \rangle \left(1 + \frac{\sigma^2_{L_{HK}}}{\langle L_{HK} \rangle^2} - \frac{\langle L_E L_{HK} \rangle - \langle L_E \rangle \langle L_{HK} \rangle}{\langle L_E \rangle \langle L_{HK} \rangle} + \cdots \right) \tag{E.2}$$

where $\sigma_{L_{HK}}$ is the instance-to-instance standard deviation of L_{HK}. In 2-D, the bias here is of $O(1/N)$. Without taking account of this bias, Johnson *et al.* use our scaling law — truncated to subleading order — to fit data at $N = 100$, $N = 316$ and $N = 1{,}000$, obtaining $\beta_E(2) = 0.7124 \pm 0.0001$. Fortunately, the bias affects only the subleading terms ($O(1/N)$ and beyond)

[1] The estimator of Johnson *et al.* follows a similar method used earlier by SOURLAS (1986), who measured $\overline{L_E/L_{th}}\ \overline{L_{th}}$, where L_{th} is a weighted average of kth-nearest neighbor distances for k up to 5. (This is by contrast with our own L^*, which is simply an unweighted average of k up to 2.) Since Sourlas' results were for the $d = 1$ random link case, it is difficult to compare with our simulations; let us note, however, that at $N = 100$ he found his method reduced $\sigma(K)$ by a factor of 4.

in the fit and should not affect the leading term that gives β_E (though care must be taken when using their L_E results for finite N). What they do not consider, however, is a test of goodness-of-fit; instead they estimate the $\beta_E(2)$ error bar "conservatively" by adding up the error bars for the three data points in the fit.[2]

Let us attempt to perform the analysis somewhat more carefully, in order to provide a more meaningful comparison between their conclusions and ours. As L_E and L_{HK} clearly are correlated very closely, we shall make the assumption that the bias in (E.2) can in fact be neglected — at least at $N = 316$ and $N = 1{,}000$ — and simply combine their data at these two points with our own data at smaller N. Table E–1 sumarizes the effects of this, comparing the coefficients in the scaling series $\langle L_E(N,2)\rangle = \sqrt{N}\,\beta_E(2)\,[1 + A(2)/N + B(2)/N^2]$, without and with these two new data points. The two results are consistent with each other: the coefficients of the fit are relatively stable (even, surprisingly, to $O(1/N^2)$), and the new χ^2 is consistent with 7 degrees of freedom (10 data points minus 3 fit parameters). The error in $\beta_E(2)$ is obtained by the standard procedure of determining the values of this quantity that make χ^2 increase by 1; those values are then $\beta_E(2)$ plus or minus one standard deviation.

Table E–1: Values of coefficients of fit $\langle L_E(N,2)\rangle = \sqrt{N}\,\beta_E(2)\,[1 + A(2)/N + B(2)/N^2]$, for previous fit ($N = \{12,13,14,15,16,17,30,100\}$ and new fit ($N = \{12,13,14,15,16,17,30,100,316,1000\}$).

Result	$\beta_E(2)$	$A(d)$	$B(d)$	χ^2
Previous fit (8 data points)	0.7120 ± 0.0002	0.1088	-1.064	5.56
New fit (10 data points)	0.7123 ± 0.0001	0.0982	-0.9819	7.97

If we stretch our assumption further and assume that the bias in (E.2) is negligible even at $N = 100$, and therefore combine our own $N = 100$ data with that of Johnson *et al.* — obtaining $\langle L_E(100,2)\rangle = 7.1286 \pm 0.0007$ — the results of "New fit" do not change at all, to the precision shown! Only χ^2 changes, decreasing now to 7.03. Another possibility is to fit exclusively to their 3 data points, though one must be cautious when using results based on so little information. Nevertheless, a two-parameter fit of the form $\sqrt{N}\,\beta_E(2)\,[1 + A(2)/N]$ gives $\beta_E(2) = 0.7123 \pm 0.0001$ and $A(2) = 0.0791$ (with $\chi^2 = 0.016$, though this value is purely anecdotal as there is only one degree of freedom!). Thus, this $\beta_E(2)$ result indeed appears credible. It is unclear why it is quoted in their article as 0.7124; however, assuming that the values quoted for the corresponding results at $d = 3$ and $d = 4$ are correct, Table E–2 compares our results and theirs.

Finally, Johnson *et al.* also performed simulations on the $d = 1$ random link case, although they did not attempt to extrapolate to $\beta_{RL}(1)$, instead noting simply that their large N data were consistent with the asymptotic value 1.0208.[3] Let us perform a fit of the usual sort on their data,

[2] Furthermore, it appears that in the notation of Johnson *et al.*, the error bar indicates the extremes of the 95% confidence interval, hence $\pm 2\sigma$. In quoting their results, we use the more standard notation of $\pm\sigma$.

[3] Their notation differs from ours by a scaling factor of 2; we cite their results following our own notation, without this factor.

Table E–2: Comparison of $\beta_E(d)$ from two different numerical studies.

d	$\beta_E(d)$ from our study	$\beta_E(d)$ from Johnson's study
2	0.7120 ± 0.0002	0.7123 ± 0.0001
3	0.6978 ± 0.0002	0.6980 ± 0.0002
4	N/A	0.7234 ± 0.0002

using the 5 points that they give: $N = 100$, $N = 316$, $N = 1{,}000$, $N = 3{,}162$, and $N = 10{,}000$. (They do not actually carry out simulations of L_{RL}/L_{HK} at $N = 10{,}000$, considering that at the level of precision used this quantity is indistinguishable from 1, based on its value of 1.000036 at $N = 3{,}162$.) The results are plotted in Figure E–1: $\beta_{RL}(1) = 1.0209 \pm 0.0002$, and $\chi^2 = 2.62$ for 2 degrees of freedom (5 data points minus 3 fit parameters). This provides excellent experimental confirmation of the cavity predictions at $d = 1$.

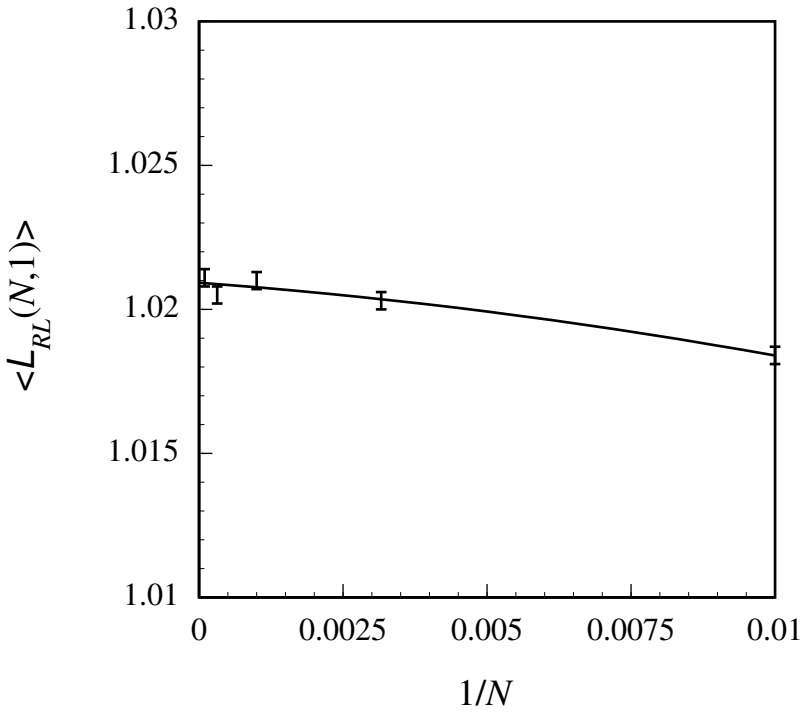

Figure E–1: Finite size scaling of $d = 1$ optimum. Best fit ($\chi^2 = 2.62$) is given by: $\langle L_{RL}(N,1)\rangle = 1.0209(1 - 0.1437/N - 10.377/N^2)$. Error bars show one standard deviation (statistical error).

General References

Armour R.S. & Wheeler J.A. (1983): "Physicist's Version of Traveling Salesman Problem: Statistical Analysis", *Am. J. Phys.*, **51**(5), pp. 405–406.

Arora S. (1997): "Polynomial Time Approximation Schemes for Euclidean TSP and Other Geometric Problems", preprint.

Arora S., Lund C., Motwani R., Sudan M. & Szegedy M. (1992): "Proof Verification and Hardness of Approximation Problems", in *Proceedings of 33rd Annual Symposium on Foundations of Computer Science*, pp. 14–23. IEEE Computer Society.

Balas E. & Toth P. (1984): "Branch and Bound Methods", in *The Traveling Salesman Problem*, ed. by Lawler E.L., Lenstra J.K., Rinnooy Kan A.H.G. & Shmoys D.B. New York: John Wiley and Sons, chap. 10.

Baskaran G., Fu Y. & Anderson P.W. (1986): "On the Statistical Mechanics of the Traveling Salesman Problem", *J. Stat. Phys.*, **45**, pp. 1–25.

Beardwood J., Halton J.H. & Hammersley J.M. (1959): "The Shortest Path through Many Points", *Proc. Cambridge Philos. Soc.*, **55**, pp. 299–327.

Bertsimas D.J. & van Ryzin G. (1990): "An Asymptotic Determination of the Minimum Spanning Tree and Minimum Matching Contents in Geometrical Probability", *Operations Res. Lett.*, **9**, pp. 223–231.

Boutet de Monvel J.H. (1996): "Physique Statistique et Modèles à Liens Aléatoires", Ph.D. thesis, Université Paris-Sud, Orsay, France.

Brunetti R., Krauth W., Mézard M. & Parisi G. (1991): "Extensive Numerical Solutions of Weighted Matchings: Total Length and Distribution of Links in the Optimal Solution", *Europhys. Lett.*, **14**(4), pp. 295–301.

Cerf N.J., Boutet de Monvel J., Bohigas O., Martin O.C. & Percus A.G. (1997): "The Random Link Approximation for the Euclidean Traveling Salesman Problem", *J. Phys. I France*, **7**(1), pp. 117–136.

Černy V. (1985): "Thermodynamical Approach to the Traveling Salesman Problem: An Efficient Simulation Algorithm", *J. Optimization Theory Appl.*, **45**, p. 41.

Cheeger J., Muller W. & Schrader R. (1984): "On the Curvature of Piecewise Flat Spaces", *Comm. Math. Phys.*, **92**(3), pp. 405–454.

Cook S.A. (1971): "The Complexity of Theorem-Proving Procedures", in *3rd Annual ACM Symposium on Theory of Computing*, pp. 151–158.

De Gennes P.G. (1972): "Exponents for the Excluded Volume Problem as Derived by the Wilson Method", *Phys. Lett. A*, **38**(5), pp. 339–340.

Efron B. & Stein C. (1981): "The Jackknife Estimate of Variance", *Ann. Statist.*, **9**, pp. 586–596.

Eisenhart L.P. (1909): *A treatise on the differential geometry of curves and surfaces*. Boston: Ginn and Company.

Fiechter C.N. (1994): "A Parallel Tabu Search Algorithm for Large Traveling Salesman Problems", *Discrete Appl. Math.*, **1**, pp. 243–267.

Garey M.R. & Johnson D.S. (1979): *Computers and Intractability: A Guide to the Theory of NP-Completeness*. New York: Freeman.

Held M. & Karp R.M. (1970): "The Traveling Salesman Problem and Minimum Spanning Trees", *Operations Res.*, **18**, pp. 1138–1162.

Itzykson C. & Drouffe J. (1989): *Statistical Field Theory*. Cambridge: Cambridge University Press.

Jaillet P. (1993): "Cube Versus Torus Models and the Euclidean Minimum Spanning Tree Constant", *Ann. Appl. Probability*, **3**(2), pp. 582–592.

Johnson D.S. (1990): "Local Optimization and the Traveling Salesman Problem", in *Proceedings of the 17th Colloquium on Automata, Languages, and Programming*. Berlin: Springer-Verlag, pp. 446–461.

Johnson D.S. & McGeoch L.A. (1997): "The Traveling Salesman Problem: A Case Study in Local Optimization", in *Local Search in Combinatorial Optimization*, ed. by Aarts E.H.L. & Lenstra J.K. New York: John Wiley and Sons, to appear.

Johnson D.S., McGeoch L.A. & Rothberg E.E. (1996): "Asymptotic Experimental Analysis for the Held-Karp Traveling Salesman Bound", in *7th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 341–350, Atlanta, GA.

Johnson D.S. & Papadimitriou C.H. (1985): "Computational Complexity", in *The Traveling Salesman Problem*, ed. by Lawler E.L., Lenstra J.K., Rinnooy Kan A.H.G. & Shmoys D.B. New York: John Wiley and Sons, chap. 5.

Karp R.M. (1977): "Probabilistic Analysis of Partitioning Algorithms for the Traveling-Salesman in the Plane", *Math. Oper. Res.*, **2**, pp. 209–224.

KARP R.M. & STEELE M. (1985): "Probability Analysis of Heuristics", in *The Traveling Salesman Problem*, ed. by Lawler E.L., Lenstra J.K., Rinnooy Kan A.H.G. & Shmoys D.B. New York: John Wiley and Sons, chap. 6.

KIRKPATRICK S. (1984): "Optimization by Simulated Annealing: Quantitative Studies", *J. Stat. Phys.*, **34**, pp. 975–986.

KIRKPATRICK S. & TOULOUSE G. (1985): "Configuration Space Analysis of Travelling Salesman Problem", *J. Phys. France*, **46**, pp. 1277–1292.

KRAUTH W. (1989): "Physique Statistique des Réseaux de Neurones et de l'Optimisation Combinatoire", Ph.D. thesis, Université Paris-Sud, Orsay, France.

KRAUTH W. & MÉZARD M. (1989): "The Cavity Method and the Travelling-Salesman Problem", *Europhys. Lett.*, **8**, pp. 213–218.

LEE J. & CHOI M.Y. (1994): "Optimization by Multicanonical Annealing and the Traveling Salesman Problem", *Phys. Rev. E*, **50**, pp. R651–R654.

LIN S. (1965): "Computer Solutions of the Traveling Salesman Problem", *Bell Syst. Tech. J.*, **44**, pp. 2245–2269.

LIN S. & KERNIGHAN B. (1973): "An Effective Heuristic Algorithm for the Traveling Salesman Problem", *Operations Res.*, **21**, pp. 498–516.

MARTIN O., OTTO S.W. & FELTEN E.W. (1991): "Large-Step Markov Chains for the Traveling Salesman Problem", *Complex Systems*, **5**(3), pp. 299–326.

——— (1992): "Large-Step Markov Chains for the TSP Incorporating Local Search Heuristics", *Operations Res. Lett.*, **11**, pp. 219–224.

MARTIN O.C. & OTTO S.W. (1996): "Combining Simulated Annealing with Local Search Heuristics", *Ann. Operations Res.*, **63**, pp. 57–75.

MÉZARD M. & PARISI G. (1985): "Replicas and Optimization", *J. Phys. Lett. France*, **46**, pp. L771–L778.

——— (1986*a*): "A Replica Analysis of the Travelling Salesman Problem", *J. Phys. France*, **47**, pp. 1285–1296.

——— (1986*b*): "Mean-Field Equations for the Matching and the Travelling Salesman Problems", *Europhys. Lett.*, **2**, pp. 913–918.

MÉZARD M., PARISI G., SOURLAS N., TOULOUSE G. & VIRASORO M. (1984): "Replica Symmetry Breaking and the Nature of the Spin Glass Phase", *J. Phys. France*, **45**, pp. 843–854.

MÉZARD M., PARISI G. & VIRASORO M.A. (1986): "SK Model: the Replica Solution without Replicas", *Europhys. Lett.*, **1**, pp. 77–82.

——— (eds.) (1987): *Spin Glass Theory and Beyond*. Singapore: World Scientific.

Nakahara M. (1990): *Geometry, Topology and Physics*. Bristol: A. Holger.

Ong H.L. & Huang H.C. (1989): "Asymptotic Expected Performance of Some TSP Heuristics: An Experimental Evaluation", *Eur. J. Operational Res.*, **43**, pp. 231–238.

Orland H. (1985): "Mean-field Theory for Optimization Problems", *J. Phys. Lett. France*, **46**, pp. L763–L770.

Padberg M.W. & Rinaldi G. (1987): "Optimization of a 532-City Symmetric Traveling Salesman Problem by Branch and Cut", *Operations Res. Lett.*, **6**(1), pp. 1–7.

Papadimitriou C.H. (1977): "The Euclidean TSP is NP-Complete", *Theor. Comput. Sci.*, **4**(3), pp. 237–244.

Papadimitriou C.H. & Steiglitz K. (1982): *Combinatorial Optimization: Algorithms and Complexity*. Englewood Cliffs, NJ: Prentice Hall.

Percus A.G. & Martin O.C. (1996): "Finite Size and Dimensional Dependence in the Euclidean Traveling Salesman Problem", *Phys. Rev. Lett.*, **76**, pp. 1188–1191.

Regge T. (1961): "General Relativity without Coordinates", *Nuovo Cimento*, **19**, pp. 558–571.

Reinelt G. (1991): "TSPLIB — A Traveling Salesman Problem Library", *ORSA J. Comput.*, **3**, pp. 376–384.

Rhee W.T. & Talagrand M. (1989): "A Sharp Deviation Inequality for the Stochastic Traveling Salesman Problem", *Ann. Probability*, **17**(1), pp. 1–8.

Sherrington D. & Kirkpatrick S. (1975): "Solvable Model of a Spin-Glass", *Phys. Rev. Lett.*, **35**(26), pp. 1792–1796.

Smith W.D. (1989): "Studies in Computational Geometry Motivated by Mesh Generation", Ph.D. thesis, Princeton University, Princeton, NJ.

Sourlas N. (1986): "Statistical Mechanics and the Travelling Salesman Problem", *Europhys. Lett.*, **2**, pp. 919–923.

Steele M. (1981): "Subadditive Euclidean Functionals and Nonlinear Growth in Geometric Probability", *Ann. Probability*, **9**(3), pp. 365–376.

Stein D. (1977): "Scheduling Dial-a-Ride Transportation Systems: An Asymptotic Approach", Ph.D. thesis, Harvard University, Cambridge, MA.

Vannimenus J. & Mézard M. (1984): "On the Statistical Mechanics of Optimization Problems of the Travelling Salesman Type", *J. Phys. Lett. France*, **45**, pp. L1145–L1153.

Weinberg S. (1972): *Gravitation and Cosmology: principles and applications of the general theory of relativity*. New York: Wiley.